Université d'Ottawa · University of Ottawa

Performance-Oriented View Planning for Automated Object Reconstruction

William R. Scott

A Thesis Submitted to
the School of Graduate Studies and Research
University of Ottawa
in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy in Electrical Engineering

Ottawa-Carleton Institute for
Electrical and Computer Engineering
School of Information Technology and Engineering
Faculty of Engineering
University of Ottawa
Ottawa, Ontario, Canada

0-612-67223-9

To Andrée

Abstract

There is a growing demand for 3D virtual models of complex physical objects in a wide range of applications. Thus far, the task of reconstructing high quality object models with range cameras (a shape measurement sensor) has been a manual, time consuming process. There is a growing demand to automate the model building process. The key open problem is to find an accurate, robust and efficient view planning algorithm - that is, the process of determining a suitable set of sensor viewpoints and imaging parameters for a specified reconstruction task with a given imaging environment.

This thesis presents and analyzes the performance of a model-based approach to view planning for automated object reconstruction. The view planning approach is "performance-oriented" in the sense that it is based on a set of explicit quality requirements expressed in a model specification and incorporates performance models for the range camera and positioning system. A new theoretical framework for the view planning problem is presented as an instance of the set covering problem with a registration constraint and is formulated as an integer programming problem. The theoretical framework facilitates a more intuitive understanding of the nature of view planning. It also provides an explicit mathematical formulation amenable to a well understood set of exact and approximate solution techniques. The measurability impact of pose error is studied in detail and counter-measures are presented to mitigate its effects. Methods for sparse sampling of object surface space and viewpoint space are presented and their performance is characterized.

Acknowledgments

I am indebted to many people for making graduate studies such a positive learning experience. First and foremost, the love and support of my dear wife Andrée kept me going during the inevitable difficult moments and encouraged me to achieve my dream. As always, my children David, Ian and Michelle were a source of great pride and inspiration. My parents, Ralph and Elizabeth, planted the seeds for this adventure by nurturing an inquisitive frame of mind from my earliest childhood days.

I conducted most of my research at the laboratories of the National Research Council (NRC), Institute for Information Technology (IIT), Visual Information Technology Group (VIT) in Ottawa with the kind permission and support of Jacques Domey, Group Leader, and Marc Rioux, Principal Research Officer. Many of my NRC colleagues helped in my work and I learned much about high precision, high definition 3D imaging from them. They are too numerous to mention everyone but particular thanks are due to Angelo Beraldin, Francois Blais, Pierre Boulanger, Luc Cournoyer, Shadia Elgazzar, Guy Godin, Michael Greenspan and Chang Shu. It was an honour to have worked with such distinguished scientists. Pierre Gariepy and Doug Taylor gave invaluable system administrative support. Dan Gamache provided excellent video recording and processing service while Anita Mainville took care of administrative issues with efficiency and good humour.

Early in my doctoral studies, the support of Dr. E. Petriu, Director of the School of Information Technology and Engineering (SITE), University of Ottawa, was instrumental in sparking my interest in computer vision and in getting my program of studies off the ground. I appreciate his faith in me and his support. Throughout my studies, Mme. Lucette Lepage handled graduate studies administration in a most pleasant and efficient way.

Particular thanks are due to my co-supervisors, Dr. J.-F. Rivest of the University of Ottawa SITE, and Dr. G. Roth of NRC/IIT. I owe much to their guidance and encouragement. Thanks also the members of my thesis committee: Prof. P. Allen of Columbia University, Prof. V. Aitken of Carleton University and Prof. R. Laganiere and Prof. P. Payeur from the University of Ottawa.

Last but not least, I wish to express my appreciation for scholarship support from the National Science and Engineering Research Council (NSERC), Ontario Graduate Scholarships program (OGS), and the University of Ottawa.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

This chapter provides an overview of the problem, defines the thesis scope, defines view planning requirements, proposes some performance measures, describes the research contributions and outlines the thesis document.

1.1 Problem Overview

1.1.1 Motivation

There is a growing demand for high quality 3D virtual models of complex physical objects in a wide range of applications, including industrial, training, medical, entertainment, cultural, architectural and others. Computer graphics can produce synthetic models for some of these applications by a combination of artistic and technical design processes. In the computer vision field, on the other hand, model acquisition is achieved by measuring the shape and visual properties of real physical objects. Numerous applications require the computer vision approach to object surface reconstruction.

Thus far, the task of object shape reconstruction with geometric measurement sensors has been a manual, time consuming process. The associated hardware and software infrastructure is costly. Details of the reconstruction process are complex and subtle. Education, training, experience and skill requirements for technical personnel are high. Consequently, there is a growing imperative to automate or semi-automate the object reconstruction process. While a limited number of commercial systems [Hym], [Dig] offer semi-automated tools to assist the reconstruction technician, there is presently no general purpose commercial system for automated object reconstruction. The principal benefits of enhanced view planning technology for automated object reconstruction will be improved productivity, improved quality, faster throughput, reduced operator skill level requirements, reduced rework requirements and lower cost.

Two broad classes of shape measuring sensors are used for 3D object reconstruction - mechanical and optical. Mechanical Coordinate Measuring Machines (CMMs) use precise mechanical movement stages and touch probes to achieve very high measurement precision - around $1\mu m$ for a high end machine that is suitably calibrated and operated by a skilled technician. However, such precision comes with high acquisition and operating costs. As a result of their relatively low data acquisition rate, CMMs are generally suitable for sparse sampling applications only. While measurement precision of the best passive or active optical range cameras is an order of magnitude inferior to that of CMMs, their capital and operating costs are substantially lower and they are capable of dense sampling at high throughput rates.

1.1.2 Scope

This thesis deals with the problem of view planning [1] for automated high quality object reconstruction by means of active, triangulation-based range

[1] Clarifications on the scope of the thesis are contained in Chapter 3 on model-based performance-oriented view planning.

cameras. In this context, view planning is the process of determining a suitable set of viewpoints and associated imaging parameters for a specified object reconstruction task with a range camera and positioning system. By object reconstruction we mean acquisition of a virtual computer model of the surface of a physical object. This is normally in the form of a triangulated polygonal mesh representation of surface geometry. Alternative surface representations such as splines, swept cylinders and super quadratics are advantageous in special cases but are neither as general purpose nor as flexible as a triangular mesh. While the focus of the work is on object reconstruction, techniques developed here are also generally applicable to inspection. The latter is an easier problem as it starts with a defined CAD model.

1.1.3 Imaging Environment

The imaging environment (Figure 1.1) for object reconstruction consists of a range camera, positioning system, various fixtures and, of course, the object.

Range Camera The principal component of the imaging environment is a range camera - a sensor for 3D shape measurement. A wide variety of technologies are used for measuring object shape. See, for example, the classic review at [Bes89]. Range cameras can be categorized into active and passive devices. The most common passive range-finding technology is stereo vision, a technique humans are intimately familiar with. Passive stereo can provide accurate measurement at short stand-off distances. In general, passive stereo depends on visual texture on the target object, is subject to illumination constraints and interference, and can provide only sparse depth measurement. With their own integral illumination source, active sensors are capable of dense depth measurement and are far less susceptible to external interfer-

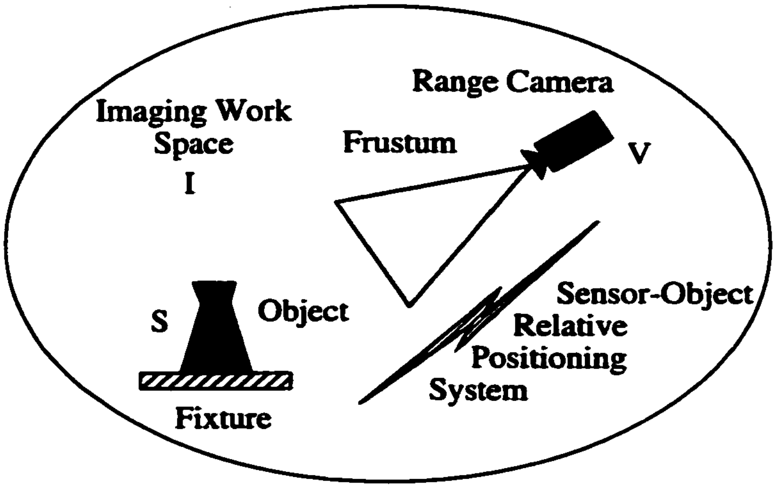


Figure 1.1: Imaging Environment

ence. However, the active illumination source (frequently a laser) may impose safety constraints on system use. Active range sensors can be divided into two sub-categories, time-of-flight and triangulation. Time-of-flight systems require very accurate timing sources and typically provide modest resolution (a few centimeters) for longer range applications i.e. tens to thousands of meters. Many different types of triangulation sensors have been developed and are in wide use. All are based on the same principle of triangulating a measurement spot on the object from a physically-separate camera optical source and detector. In general, active triangulation-based range cameras are capable of very precise ($100\mu m$ or better), dense depth measurements (many samples per square millimeter) over relatively small sensor frustums i.e. up to about a meter in stand-off distance.

The focus of the thesis work is on triangulation-based active laser range scanners as they are widely used for precise measurement over relatively small imaging workspaces for both object reconstruction and inspection. Several of their attributes are important for view planning. The optical baseline, the distance between the laser and the optical receiver, is significant with respect to the measurement stand-off distance to the object. Consequently, coverage shadow zones[2] arise with respect to visibility by the laser, the optical receiver or both. The sensor field of view (FOV), depth of field (DOF) and therefore frustum volume are limited. Measurement precision and sampling density are highly non-uniform within the frustum. Additionally, the measurement process is subject to random non-isotropic geometric noise and several artifact phenomena.

Positioning System Imaging all sides of a three-dimensional object requires a variety of viewing perspectives. Thus, a positioning system is required to move the sensor, the object or both in some combination to achieve the desired viewing perspectives. An imperfect positioning system introduces pose error which has two consequences. Firstly, a planned view is not the one actually executed. Sufficiently severe pose error can render view planning futile. Secondly, the resulting pose uncertainty necessitates an image registration stage.

Viewpoint Space Viewpoints should be treated as generalized viewpoints [TTA95], [TAT95], [RM97], [Pit97a], consisting of sensor pose and a set of controllable sensor parameters. Thus each viewpoint has an associated sensor configuration. Sensor pose is limited by the degrees of freedom and range of motion of the positioning system. Viewpoint space V is the set of generalized viewpoints defined by the range and sampling of these parameters.

[2]Typical shadow effects can be seen in Figure 4.19.

Object The object is the feature to be modeled. Specifically, we wish to capture its 3D shape. Object surface space S is the set of 3D surface points sampled on the object. These can be considered as vertices in the resulting object mesh model.

Fixtures Gravity requires that the object be supported by a fixture of some kind. Consequently, acquisition of all-aspect views will require the object to be repositioned on the fixture at least once and more likely several times. After each repositioning, the new relative pose must be determined to bring the system into a common coordinate frame. Fixtures, the positioning system and any other structures in the imaging workspace I introduce occlusion and collision avoidance considerations.

1.1.4 Reconstruction Cycle

The classical object reconstruction or model building cycle (Figure 1.2) consists of four main phases - plan, scan, register and integrate. A sequence of views must be planned to derive the *view plan* or *next-best-view (NBV) list* N. The sensor must be translated and rotated to the appropriate pose and configured with appropriate settings, after which the physical scanning process must be undertaken. Unless the positioning system is error-free, acquired range images must be registered by means of an image-based registration technique such as the standard Iterative Closest Point (ICP) method [BM92] to bring them into a common coordinate frame. Subsequently, registered images must be combined into a single, non-redundant model - a process commonly known as integration. The process continues until some stopping criteria are satisfied.

These are the basic reconstruction steps. Other related activities are

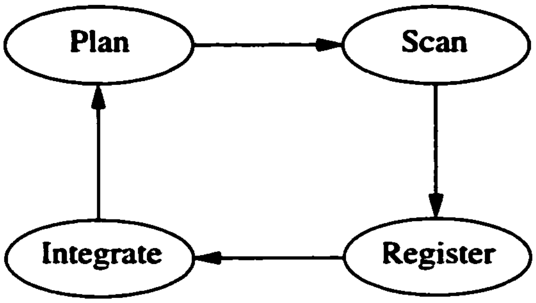


Figure 1.2: Object Reconstruction Cycle

briefly as follows. Calibration of both sensor and positioning system is an essential prerequisite and may need to be repeated after certain positioning system reconfigurations. Filtering of noise and artifacts is required at several stages in the model building process. Additionally, depending on the application, there may be a requirement for model compression or decimation and for texture mapping of reflectance data onto the geometry.

While challenges remain, adequate solutions exist for the scan-register-integrate portions of the process to satisfy the goals of automated object reconstruction[3]. On the other hand, the view planning component remains an open problem despite two decades of research. For large reconstruction tasks, the image registration and integration (model building) tasks can be very time consuming. By the time deficiencies in the partial model are discovered, the imaging team may have left the site or access to the object may no longer be readily available. Hence, there is a need for a view planning scheme capable of developing reliable view plans in a timely manner for

[3]Research on ICP refinements remains an active field of research. For example, see [RL01] and [LGG01]. Research on model building also remains active, with a number of open issues [Rot00].

comprehensive object coverage in accordance with all technical criteria of the reconstruction task.

1.1.5 The View Planning Problem

From the foregoing, it is apparent that the view planning problem (VPP) involves reasoning about the state of knowledge of three spaces - object surface space S, viewpoint space V and imaging workspace I. While tied to three-dimensional surface geometry, S is amenable to a 2D parameterization. In the unrestricted case, the pose component of viewpoint space V is six dimensional - three position and three rotation. Configurable sensor parameters such as laser power and scan length can further raise the dimensionality of generalized viewpoint space. Thus, we indicate the dimensionality of viewpoint space as $6D^+$. Imaging workspace I is the 3D workspace capable of being viewed by the sensor over the full range of the positioning system. It is of concern primarily for visibility analysis and collision avoidance considerations. The complexity of the VPP is immediately apparent from the high dimensionality of the search spaces in S, V and I.

This work concerns view planning for *performance-oriented* reconstruction [SRR00] which is defined as model acquisition based on a set of explicit quality requirements expressed in a *model specification*. In addition to all-aspect coverage, measurement quality is specified in terms of precision, sampling density and perhaps other quality factors. Performance-oriented view planning requires suitable models of both sensor and positioning system performance. Specifically, it requires:

- a *sensor model* to include a description of the camera geometry, frustum geometry and characterization of measurement performance within the calibrated region, and

- a *positioning system model* describing the degrees of freedom, range of motion and positioning performance within the movement envelope.

Given this imaging environment and reconstruction cycle, the view planning task is deceptively simple, yet computationally complex. Expressed informally, the view planning probiem is as follows:

> For a given imaging environment and target object, find a suitably short view plan N satisfying the specified reconstruction goals and achieve this within an acceptable computation time.

N will be a subset (preferably a very small subset) of viewpoint space V i.e. $N \subset V$. We will shortly see why it is impractical in most realistic reconstruction cases to seek an absolute minimal length view plan. What is an "acceptably short" computation time is application dependent. In an industrial setting, the object reconstruction application involves repetitive modeling of different objects. In such a case, throughput is critical so the emphasis is on time-efficient algorithms, perhaps trading off some aspects of quality, such as the length of the view plan, in the interests of time efficiency. In some reconstruction tasks, the quality of the model is the primary objective and time efficiency will be secondary. Inspection applications involving the repetitive execution of an inspection plan on a production line will place a premium on view plan efficiency over the time taken to create it.

1.2 View Planning Requirements

In this section, we define requirements for the view planning process in greater detail[4]. The requirements are stated from a research perspective. Additional

---

[4]Pito [Pit97a] provides a good description of the view planning problem and defines constraints on the choice of the next-best-view and desirable properties of view planning

development and system integration details will arise when moving prototype view planning algorithms to a production configuration of an automated object reconstruction system.

1.2.1 Assumptions

To begin, our definition of the view planning problem for object reconstruction is based on several assumptions:

- Both the range camera and positioning system are calibrated and calibration parameters are available to the view planning module.
- The imaging workspace is compatible with object size and shape.
- Range camera performance is compatible with the model specification for the reconstruction task.
- The sensor and positioning system have compatible performance. Specifically, positioning system pose error is compatible with the volume of the range camera frustum.
- Object shape is compatible with sensor capabilities.
- Object material properties are compatible with sensor capabilities.

The last two items merit elaboration. Range cameras are inherently sampling devices. Therefore, the reconstruction process is subject to the limitations of a sampled representation of continuous surface shape. The familiar sampling theorem applicable to sampled signals in one and two dimensions does not extend to three dimensions. As there is no natural low pass filter for 3D physical shape, aliasing effects can be anticipated. Convoluted shapes

---

algorithms.

can easily be conceived which cannot be fully imaged by a range camera, for example a gourd or a sea urchin-like shape. Deep cavities, deep holes or multiple protrusions will be troublesome or completely impractical to image. Therefore, practicality requires that object geometry and topology be "reasonable".

Laser scanning sensors experience difficulty with certain materials. They perform best with material having surface scattering properties and a relatively uniform reflectance, neither overly absorbent nor overly reflective. Natural or man-made material with volume scattering properties such as hair and glass produce multiple reflections over a surface depth zone, degrading or prohibiting accurate range measurement.

The dynamic range of the current generation of range sensors is limited. Excessively absorbent material will scatter insufficient light, causing "drop outs" i.e. missing data. Shiny surfaces are characterized by specular reflection. Depending on illumination and observation geometry, specular surfaces may result in drop outs due to insufficient scattered energy or outliers or drop outs (depending on camera design) due to receiver saturation from excessive received energy. Multiple reflections on a shiny surface in corner regions can also produce wild measurements. Finally, abrupt reflectance changes will produce edge effects in the form of measurement biases and artifacts [CL96].

Consequently, the current work is limited to view planning for objects with "reasonable" shape and benign reflectance characteristics. The issue of extending view planning to objects with more difficult reflectance characteristics is addressed in the concluding chapter under future work.

Category	Requirement
General	Model quality specification
	Generalizable algorithm
	Generalized viewpoints
	View overlap
	Robust
	Efficient
	Self-terminating
Object	Minimal a priori knowledge
	Shape constraints
	Material constraints
Sensor	Frustum
	Shadow effect
	Measurement performance
Positioning System	6D pose
	Pose constraints
	Positioning performance

Table 1.1: View Planning Requirements

1.2.2 Constraints and Requirements

The following constraints and requirements apply to the view planning task for object reconstruction. They are summarized in Table 1.1 under the categories of general, object, sensor or positioning system. The following amounts to a "research specification" for a view planning system. Where objectives are quantified, it is presumed that research-quality algorithms are subject to refinement, development and system integration on suitable production hardware and software platforms.

Model Quality Specification By definition, performance-oriented view planning is based on a model specification containing explicit, quantified model quality requirements. In our work, we specify measurement precision and sampling density. Additionally, there is an implicit requirement for 100% surface coverage. In other applications such as inspection, it may be appropriate to limit coverage to specified object regions.

Generalizable Algorithm For wide applicability, the view planning technique should apply to a broad class of range sensors, positioning systems and objects, while staying within our assumptions.

Generalized Viewpoints In addition to sensor pose (position and orientation), the view planning algorithm should be capable of planning reconfigurable sensor parameters such as laser power and scan length through the use of generalized viewpoints. A generalized viewpoint $(\mathbf{v}, \lambda_s)$ consists of sensor pose $\mathbf{v}$ and a set of controllable sensor parameters λ_s.

View Overlap The view planning algorithm should provide the degree of image overlap necessary for integration and registration. Image integration requires image overlap along the boundaries of adjacent images in the view plan. If pose error is such that image-based registration is required, the image set should have sufficient overlap and sufficient shape complexity in overlapping regions for registration of the image set.

Robust The view planning algorithm should be immune to catastrophic failure, handle sensor and positioning system noise and artifacts and require minimal operator intervention.

Efficient The view planning algorithm should be efficient enough to be competitive with skilled human operators. As an initial goal, the algorithm should be capable, with production-quality hardware and software, of producing a specification-compliant view plan for a moderately complex object within one hour.

Self-terminating The view planning algorithm should recognize when the goal has been achieved or when progress towards it has stalled.

Limited a priori knowledge The view planning algorithm should be effective with minimal a priori object knowledge, specifically no more than approximate bounding dimensions and centroid.

Shape constraints The view planning algorithm should be effective with all reasonable object geometry and topology that is compatible with sensor measurement capabilities.

Material constraints The view planning algorithm should be effective with all object material properties that are compatible with sensor measurement capabilities.

Frustum The shape of the viewing frustum should be modeled. This includes sensor depth of field (DOF), field of view (FOV) and scan length or scan arc.

Shadow effect The bi-static nature of the sensor should be modeled - that is, the physical separation of the laser and detector.

Measurement performance Sensor measurement performance should be modeled. This includes variation of measurement precision and sampling density within the frustum and surface inclination effects. The following range camera artifacts should be modeled: geometric step edges, reflectance step edges and multiple reflections.

6D pose The view planning algorithm should handle a positioning system with an unconstrained pose space, three position and three rotation.

Pose constraints The view planning algorithm should model constraints on the degrees of freedom and range of motion of the positioning system.

Positioning system performance Positioning system performance should be modeled, to include pose error and repositioning/reconfiguration time.

Capabilities of current view planning techniques will be examined with respect to these requirements and issues in the next chapter.

1.3 Performance Measures

The following measures are proposed for evaluating view planning algorithm performance for a given reconstruction task with a given imaging environment:

- *View plan quality* The quality of a view plan is determined by the quality of the reconstruction it generates. This can be expressed as the overall verified measurability[5] m_v of the reconstruction with respect to the model specification. m_v is an indicator of the fidelity and robustness

[5]Measurability is defined in Chapter 3.

of the view planning process plus the adequacy of the discretization schemes for viewpoint space and surface space. The goal is $m_v = 1.0$.

- *View plan efficiency* A simple measure of view plan efficiency is the length of the generated view plan relative to the optimum achievable for that task i.e. $e_v = n_{Opt}/n$, where $n = |N|$. As it may be impractical to determine with certainty the optimum view plan length n_{Opt} for complex reconstruction tasks, a suitable surrogate is the length of the best solution n_{Best} found thus far amongst all view planning techniques examined for the same task i.e. $n_{Opt} \approx n_{Best}$. e_v is an indicator of the efficiency and completeness of the discretization schemes for viewpoint and surface space and the efficiency of the set covering process. The goal is $e_v = 1.0$.

- *View plan computational efficiency* This factor evaluates the computational cost of generating the view plan. Two performance measures are computational complexity and execution time on a defined platform. Each of these measures has weaknesses, but together they provide a reasonable appreciation of algorithm performance. Computational efficiency indicates the efficiency of the discretization schemes for viewpoint and surface space, the efficiency of visibility analysis and the efficiency of the set covering process. Computational efficiency goals are application-dependent.

1.4 Contributions

The main contributions of this thesis are summarized as follows:

- Development of a new theoretical framework for the view planning problem (VPP) as an instance of the set covering problem (SCP) with a registration constraint [SRR01a], [SRR01c].
- Development, implementation and characterization of a model-based performance-oriented view planning concept [SRR00]. Performance-oriented view planning is defined as view planning with respect to objective, quantified technical requirements specified in a model specification.
- Characterization of sparse and approximate sampling of object surface space [SRR01b].
- Development and characterization of model-based viewpoint generation algorithms for efficient, sparse sampling of viewpoint space.
- Characterization of the performance impact of pose error due to positioning system imperfections and identification of counter-measures to mitigate the effects of pose error [SRR01d], [SRR01e], [Sco02].

Prior theoretical treatment of the view planning problem has been limited. The new theoretical framework presented in this thesis makes clear the isomorphism between the view planning problem and the well known set covering problem of combinatorial mathematics and therefore explains why the VPP is NP-complete. In addition to providing a basis for a more intuitive understanding of the nature of the problem, this framework provides an explicit solution formulation as an integer programming problem while addressing image registration requirements. A large body of techniques and experience exists for integer programming and solving the SCP, including both exact and approximate solution techniques. We show that a constrained greedy search set covering algorithm is appropriate for view planning in object reconstruction.

The performance-oriented model-based view planning concept (Modified Measurability Matrix or 3M) incorporates a number of innovations: an explicit model quality specification, more detailed sensor and positioning system models than heretofore used, objective rather than subjective success criteria and techniques for sparse sampling of object surface space and viewpoint space. The 3M technique is a readily parallelizable algorithm.

The characterization of sparse sampling of object surface space provides some engineering rules of thumb for sampling density and surface noise tolerance in the exploratory model. The parameters examined are sampling density, surface noise and object shape.

Efficient techniques for sparse sampling of viewpoint space are developed and characterized. These effectively reduce the pose component of viewpoint space from six dimensions to two or three, depending on the algorithm. Viewpoints are treated as generalized viewpoints and reconfigurable sensor parameters are optimized on a per-viewpoint basis. The sampling of viewpoint space is not fixed a priori but is optimized on the basis of object shape and the capabilities and limitations of the sensor and positioning system.

The impact of pose error on view planning had not been examined prior to this work. The research reported herein provides a quantitative and qualitative assessment of pose error effects using a general purpose error model. Based on this analysis, a variety of countermeasures are proposed to mitigate pose error effects. The quantitative analysis should assist in setting view planning system parameters and in conducting system-level tradeoffs.

1.5 Thesis Outline

The outline of the remainder of the thesis is as follows:

- Chapter 2 reviews related work in the literature.
- Chapter 3 describes the modified measurability matrix concept and discusses computational complexity.
- Chapter 4 provides a theoretical framework for the view planning problem and presents experimental results for a greedy search set covering algorithm modified to include a registration constraint.
- Chapter 5 analyses object surface space discretization in the exploratory model, in particular sampling density and surface noise.
- Chapter 6 presents and characterizes viewpoint generation algorithms for efficient sampling of viewpoint space.
- Chapter 7 analyses the performance impact of pose error and provides view planning counter-measures.
- Chapter 8 concludes the thesis and discusses issues for future work.

Key symbols, acronyms and abbreviations are defined at Annex A.

Chapter 2

Related Work

This chapter surveys the view planning literature applicable to object reconstruction, evaluates existing methods with respect to the previously defined view planning requirements and describes the remaining open issues.

2.1 View Planning Surveys

Surveys relevant to view planning for object reconstruction can be found primarily in doctorate and masters theses published over the last decade. These are listed at Table 2.1.

In a brief and clearly written technical report [HH98] addressing only seven papers ([BYW+95], [Con85], [HK89], [MB93], [Pit96b], [TAT95], [WF90]), Hall-Holt quickly zeroes in on two key view planning issues - representation of available geometric information and how to deal with the potentially vast search space. While he does not answer the questions he poses, Hall-Holt succinctly examines from several perspectives the view planning approaches taken by some of the most notable authors. The survey concludes with a discussion of open issues: algorithm efficiency in the presence of large amounts of data, appropriate geometrical representations, hybrid models,

Year	Author(s)
1999	Flavio Prieto Ph.D thesis [Pri99]
1998	Michael Reed Ph.D thesis [Ree98]
1998	Olaf Hall-Holt Technical Report [HH98]
1997	Steven Abrams Ph.D thesis [Abr97]
1997	Brian Curless Ph.D thesis [Cur97]
1997	Nikos Massios Masters thesis [Mas97]
1997	Dimitri Papadopoulos-Orfanos Ph.D thesis [PO97]
1997	Richard Pito Ph.D thesis [Pit97a]
1997	Roberts and Marshall technical report [RM97]
1997	Yiming Ye Ph.D thesis [Ye97]
1996	Joseph Banta Masters thesis [Ban96]
1996	Leland Best Ph.D thesis [Bes96]
1996	Mark Wheeler Ph.D thesis [Whe96]
1995	Jasna Maver Ph.D thesis [Mav95]
1995	Tarabanis et al survey paper [TAT95]
1994	Stephane Aubry Ph.D thesis [Aub94]
1994	Kiriakos Kutulakos Ph.D thesis [Kut94]
1994	Chris Pudney Ph.D thesis [Pud94]
1994	Lambert Wixson Ph.D thesis [Wix94]
1993	Timothy Newman Ph.D thesis [New93]
1993	Sergio Sedas-Gersey Ph.D thesis [SG93]
1993	Glen Tarbox Ph.D thesis [Tar93]
1993	Xiaobu Yuan Ph.D thesis [Yua93]
1992	Konstantinos Tarabanis Ph.D thesis [Tar92]
1992	Besma Roui-Abidi Ph.D thesis [RA92]
1990	Michael Buzinski Masters thesis [Buz90]
1990	Seungku Yi Ph.D thesis [Yi90]

Table 2.1: View Planning Surveys

multi-resolution techniques, ray tracing avoidance, quantitative performance assessment, performance benchmarks and improved scanner models.

The most widely referenced sensor planning survey is by Tarabanis et al

[TAT95]. The authors consider three vision tasks - inspection, recognition, and reconstruction - but focus almost exclusively on the first of these. The underlying perspective is that of conventional intensity imaging, although a few references are made to range imaging. View planning is categorized as following one of three paradigms - synthesis, generate and test, or expert system. The open problems in view planning circa 1995 as seen by Tarabanis et al were:

- Modeling and incorporating other constraints such as integrating collision avoidance with view planning, dynamic sensor planning and modeling the operating range of the employed sensor.
- Modeling and incorporating other sensors such as tactile, range, force-torque and acoustic sensors.
- Relaxing some of the assumptions made in current approaches such as feature uncertainty and error characterization.
- Illumination planning to include higher order lighting and reflectance models, multiple sources, specularity and inter-reflections.
- Sensor and illumination modeling to include mapping model parameters to controllable sensor settings and accurate sensor noise models.

A comprehensive survey of automated visual inspection systems by Newman and Jain [NJ95] considers binary, intensity, colour and range image sensors for a wide range of inspection applications.

As will be seen from the surveys, view planning methods can be categorized in several different ways. From our perspective, we begin with two top level categories - model-based or non-model-based. It is convenient to then sub-categorize the two classes somewhat differently. At the end of this

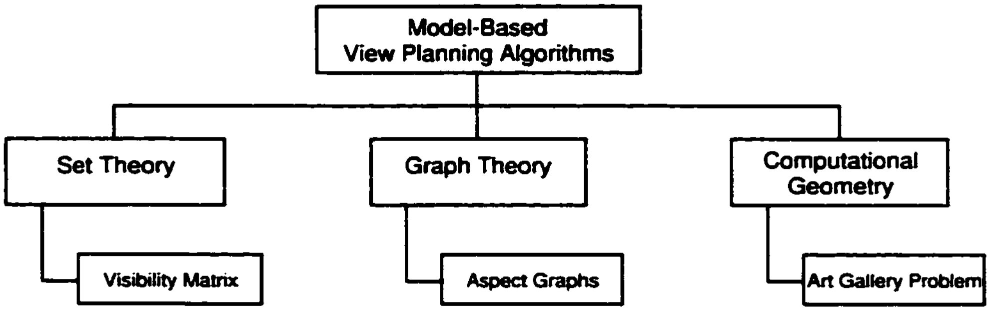


Figure 2.1: Model-Based View Planning Algorithms

chapter, current methods are evaluated with respect to the view planning requirements derived in Chapter 1.

2.2 Model-Based View Planning

Model-based view planning methods can be categorized by the representation used for the knowledge embedded in the object model (Figure 2.1).

2.2.1 Set Theory Methods

Visibility Matrices At the heart of the set theoretic approach is a visibility matrix, whose elements encode the visibility of discrete object surface points from each pose in quantized viewpoint space. All essential visibility information regarding the imaging domain is encoded in a single data structure. A more sophisticated enhancement replaces visibility by measurability, in which the data elements encode estimated measurement quality.

Tarbox and Gottschlich Tarbox and Gottschlich [TG95] incorporate elements of set theory, graph theory, computational geometry, mathematical morphology and statistical non-linear optimization in their thorough examination of view planning for automated inspection applications.

They discretize the solution space by recursive subdivision of an icosahedron and constrain solutions to lie on a viewing sphere completely containing the object with the viewing direction oriented towards the center of the sphere. They further assume that the object lies completely within the sensor's frustum. An octree-encoded voxel occupancy model is used to represent the sensing volume. Their view planning approach is tailored to a specific long-baseline, active triangulation-based range sensor so that the view planning problem becomes one of examining the set of all ordered pairs of points on the view sphere separated by the specified sensor baseline. They briefly discuss culling inadmissible points from the viewpoint set by means of ray tracing operations, such as those occluded by a mounting fixture and supporting plane. The principal consideration in viewpoint selection is taken to be the surface area that a given viewpoint is capable of sensing.

Consideration of grazing angle effects figures prominently in their approach. Different grazing angle thresholds are set for the light source and the camera. Grazing angle constraints are only considered from the perspective of their impact on dilution of surface sampling density and viewability at acute angles in the presence of surface micro-structure. This perspective masks the most important impact of grazing angle which is on measurement error. Other aspects of sensor noise and artifacts are not considered.

The authors note that it would be desirable to define the search for a good inspection plan as - "Find the smallest number of sensing operations such that the union of the subsets of surface points measurable by each operation completely covers the surface point set." However, as this problem is known to be NP-complete, the authors conclude that it will be necessary to employ a view planning algorithm that can find a satisfactory but not necessarily optimal view sequence.

Tarbox and Gottschlich develop and examine the performance of three algorithms. The first two employ an irrevocable selection strategy and differ by the manner in which grazing angle constraints are treated. The third algorithm is novel in that it employs a revocable selection strategy for viewpoint set minimization based on a randomized search with simulated annealing.

To account for slight pose errors, they remove viewpoints which are not robust to pose variation. This is accomplished by morphological operations on the viewpoint set associated with each surface point.

Their approach is based on a measurability matrix $\mathcal{M}(i,j)$ computed by a complete visibility analysis for the laser source and camera over the set of all surface points and all admissible viewpoints. Given the span of the discretized variables, the computational complexity of the approach is prohibitive. This is a fundamental limitation regarding direct applicability of the work to object reconstruction. Nevertheless, the authors' thorough and well written treatment of the subject provides useful insights into several aspects of the view planning problem.

2.2.2 Graph Theory Methods

Aspect Graphs An aspect graph of an object has a node representing every aspect of that object and arcs connecting all adjacent aspects on an object [TG95], [BD90]. An aspect is loosely defined as the set of viewpoints of the object such that the unoccluded portion of the object seen from all those viewpoints is qualitatively the same. In other words, viewpoint space is partitioned into regions providing equivalent views. Each node represents one such region while arcs represent their adjacency in viewpoint space.

While aspect graphs are an intriguing theoretical possibility, there are practical difficulties. When dealing with objects as simple polyhedra and considering a viewpoint purely in terms of the visibility of object features from that point, the criterion "qualitatively the same" simply means that the same vertices, edges and faces are visible from that viewpoint. When the object is changed from discrete to continuous, quantization of viewpoint space inherent to an aspect graph representation strictly fails. For some applications, one may redefine an aspect graph in terms of the visible topology of the object's silhouette, but this is not useful for object reconstruction. Additionally, as we are dealing with a sensing problem which involves not only visibility but mensuration, viewpoints with the same visibility (i.e. the same aspect) are by no means equivalent for sensing purposes. Secondly, aspect graphs for even moderately complex objects quickly become huge and unwieldy, as does the computational complexity of handling them.

Consequently, aspect graph theoretical development is not sufficiently mature [FMA+92] to be a suitable representation basis for performance-oriented view planning for automated object reconstruction.

2.2.3 Computational Geometry Methods

Art Gallery Problem A classic 2D computational geometry problem concerns the visibility of edges in a polygon, the so-called "art gallery problem" [Urr00], [XCB86]. Given the floor plan of an art gallery as a polygon P, the problem is to find an upper bound on the number of "guards" represented by points such that the interior walls of P are completely visible [KKK80]. The task is to determine the minimum number and placement of watchmen who collectively can see all walls in the gallery. The object reconstruction problem is somewhat related to this classic computational geometry problem. However, there are additional complexities - three dimensions, bi-static

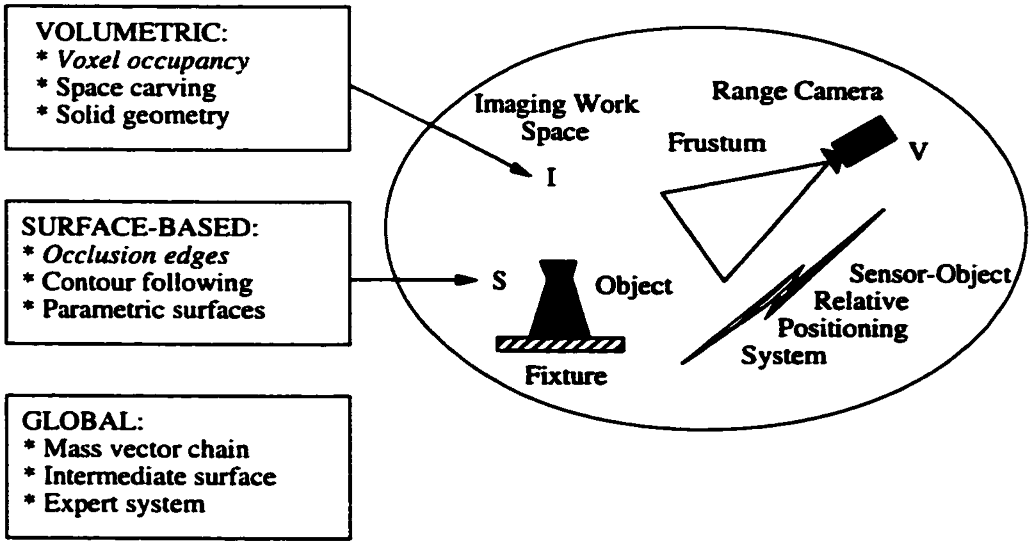


Figure 2.2: Traditional Non-Model-Based View Planning Methods

visibility (source and receiver) plus mensuration in lieu of visibility.

2.3 Non-Model-Based View Planning

It is convenient to classify existing view planning methods (most of which are non-model-based) by the domain of reasoning about viewpoints - that is, *volumetric*, *surface-based* or *global* (Figure 2.2). Some methods combine several techniques. The majority of current methods fall into two sub-categories- voxel occupancy or occlusion edges.

2.3.1 Surface-Based Methods

Surface-based algorithms reason about knowledge of object surface space S.

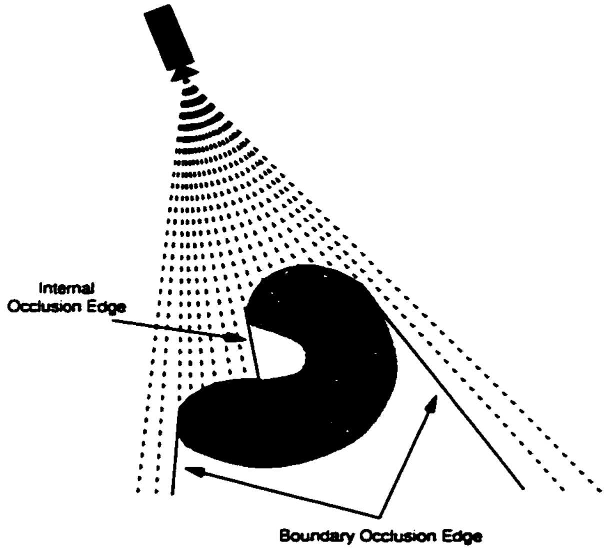


Figure 2.3: Occluding Edges in Geometric Images with Mono-static Sensor

Occlusion Edge Methods

Occlusion Edges The most commonly used traditional view planning mechanism exploits geometric jump edges. Illustrated in Figure 2.3, this approach is based on the premise that occlusion edges internal to the image indicate surface areas not yet sampled, while boundary jump edges represent the boundary of the unobserved volume. In both cases, occlusion edges provide cues for the next best viewing direction.

Maver and Bajcsy Some of the earliest papers using the occlusion edge method are by Maver and Bajcsy [MB90], [MB93]. Their approach is tailored to a long baseline laser profile scanner with a positioning system limited to a single rotational degree of freedom. The two-stage method separately considers source and receiver occlusions. The focus was on finding and labeling boundary occlusion edges. View planning is done with reference to a plane defined by the support fixture. The initial stage plans views as a rotation about the range camera boresight. The projection of occluded regions on the support plane is approximated by polygons. Occlusion-free arcs are computed by a visibility analysis making assumptions about the height of each "shadow zone pixel". A histogram is formed by summing the viewing arcs for each pixel and a next-best-view is determined from histogram maxima. The algorithm is computationally intensive due to the many-on-many visibility calculations. Performance is sensitive to the underlying assumptions.

In a second view planning stage, described analytically but not implemented, occlusion edges of source shadow zones are used in a similar manner to determine a new orientation for the camera illumination plane. The authors restrict the search for new scanning planes to the set perpendicular to the original scanning setup. Simulated results are shown.

The approach is limited by its formulation as a 2 1/2 D scene view rather than a true 3D perspective. The scene is defined as a horizontal x-y array of pixels with a height value, either measured or assumed. Coupled with sensor and positioning system limitations, the method is incapable of acquiring an all-aspect model.

Contour Following

Once having located a portion of the object, the contour following technique involves "painting" the object with the sensor by keeping it in close proximity to the surface at all times. The technique has been applied to a class of range sensors with a limited sensing volume. Collision avoidance is a primary concern. Pudney [Pud94] describes the application of contour following to a robot-mounted range sensor.

Conventional terminology labels range sensors as range "cameras", implying an output in the form of an "image". In fact, most range cameras are profile scanners which only acquire data in an image format through linear or rotational sensor motion by a fixed or programmable amount. Contour following usually involves acquiring long strips of profiles of an arbitrary length rather than a range image in the conventional sense[1]. Because the positioning systems used typically are accurate in pure translation but have limited mobility or may require recalibration in rotation, scanning is frequently done along non-linear curves with a fixed orientation. The technique is widely used commercially with mechanical CMM devices. Contour following works best with objects large in size relative to sensor coverage and for relatively simple shapes with smoothly flowing lines. It is more problematic with smaller objects and more complex shape. Other contour following work with range sensors includes [SCF98], [LBG99].

Parametric Surface Representations

Superquadric models have been widely used in machine vision applications due to their flexibility and compact representation. However, they are limited to fairly simple scenes unless the scene is segmented and piece-wise fitted with separate models. Additionally, superquadrics are highly non-linear.

Whaite and Ferrie Related more to robotic exploration than object reconstruction, the autonomous exploration approach of Whaite and Ferrie ([WF97], [WF92], [WF91], [WF90]) nevertheless provides an interesting high level perspective on characterization and exploitation of uncertainty in view planning. Their gaze-planning strategy uses model uncertainty as a basis for

[1] See also the discussion at Section 6.2.1.

selecting viewpoints. After the data is segmented, the task is to find superellipsoid parameters best describing each segmented part. They observe that the best sensor locations are those where the ability to predict is worst - that is, where variance in the fit of the data to the current model is greatest.

Because of concerns about the validity of the linearized theory, they take a conservative approach in which the sensor is always moved in small steps. Using a numerical analysis technique, they search for and move the sensor in the direction of maximum uncertainty. Movements are constrained to a view sphere at a fixed geodesic distance from the current scanner location. Experiments show the algorithm to be attracted to regions of high curvature.

The authors use a multi-layered system design. A general purpose, sensor-independent and environment-independent lower layer navigator module handles view planning. A higher level explorer module contains application-specific knowledge and handles executive level functions such as delegating tasks, monitoring progress, making decisions and resolving conflicts. While the layered design is attractive, it is difficult to see how detailed sensing and environmental factors can be separated from the act of sensor view planning.

The method directly incorporates measurements of sensor noise, although it is unclear exactly which noise sources are modeled or what the fidelity of the corresponding models is. View overlap is a consequence of the conservative nature of the search strategy rather than application of an explicit constraint.

While innovative and providing useful insights into view planning at a high level of abstraction, the technique is designed more for approximate volumetric modeling of simple scenes for robotic manipulation and recognition than for high definition surface modeling.

2.3.2 Volumetric Methods

Volumetric methods select viewpoints by reasoning about the state of knowledge of the imaging workspace I. Each scan labels a portion of I. The general strategy is to select a viewpoint with the greatest prospective reduction in uncertainty about the imaging workspace. Volumetric methods focus particularly on the solid shadows cast by the scanned object.

Voxel Occupancy Methods

Common volumetric methods involve encoding space occupancy by a voxel occupancy grid or an octree.

Voxel Occupancy Grids Voxelization is a widely used, compact means of encoding spatial occupancy. Voxel grids can also be used for a coarse surface representation, although they are clearly not suited for high precision modeling. Their principal disadvantage is the large memory requirement for even moderate volumetric quantization levels. In a typical imaging environment, the object occupies only a small portion of the imaging workspace and its surface intersects an even smaller portion of the voxelized space. Most authors ignore the impact of misalignment of spatial quantization intervals [Gre99] between views. The phenomena is similar to timing jitter in conventional time domain signal processing.

Banta et al Banta et al [BYW+95] and [BA96] define the next-best-view as “the next camera pose which will extract the greatest amount of unknown scene information”. Their objective is to minimize the number of views required for reconstruction. Some of the work is similar to Connolly's ([Con85]) although they constrain the NBV search to a smaller search space.

Both the imaging workspace and object surface are represented by voxel occupancy grids in which voxels are labelled occupied or unoccupied. Rather than allowing an "unknown" state, this binary approach labels all points not currently visible as occupied and merges views by a voxel-wise logical "AND" operation. The three papers explore several cuing mechanisms[2] for suggesting feasible views, including orienting the sensor in the direction of the viewpoint with the greatest number of potentially visible voxels based on local surface exposure[3], the viewpoint revealing the greatest number of hidden voxels based on a ray tracing visibility analysis, the mean of the three largest jump edges in the most recently acquired image or the centroid of the cluster containing the largest number of unknown voxel faces.

The notion of applying several view planning concepts in combination under intelligent control is a useful contribution. The authors also propose tests to ensure selection of a "good" NBV such as validating by ray tracing that the feature of interest is actually visible and enforcing a minimum angular separation between views on the view sphere. Additionally, they examine several termination criteria i.e. terminate when the size of either the surface model or the occluded model ceases to change by a significant amount or when the ratio of the size of the surface model to that of the occluded model is "large". However, the proposed termination criteria do not relate to objective performance requirements for the target model.

The various algorithms attempted were effective in what could be called exploratory, approximate view planning for topologically complex objects with purely synthetic data but were less effective in acquiring smaller geometric detail. While innovative, the work is subject to a number of sim-

[2]The authors also discuss a method based on locating surface regions of maximal curvature which they appear to incorrectly associate with geometric jump edges.

[3]This is similar to Connolly's "normal" algorithm which fails in concave regions.

plifications and limitations. The range camera is treated as an error-free mono-static sensor with uniform performance within the imaging workspace. Pose error is ignored. Viewpoints are restricted to a fixed radius view sphere.

Massios and Fisher A paper [MF98] summarizing Nikos Massios' Master's thesis [Mas97] builds on previous voxel occupancy methods by using the weighted sum of visibility and "quality" factors as a NBV objective function.

$$f_{total}(\vec{v}) = w_v f_{visiblity}(\vec{v}) + w_q f_{quality}(\vec{v})$$

The visibility cue is taken from an occlusion edge analysis finding occlusion plane voxels which are defined as unseen voxels with an empty neighbour voxel. Visibility factor $f_{visiblity}$ is set to the number of occlusion plane voxels visible from viewing direction $\vec{v}$, as determined by ray tracing. The absolute value of the inner product of the estimated local surface normal and the viewing direction vector is taken as the local quality measure for occupied voxels. A region quality estimate is also formulated for each pixel due to unspecified system inaccuracies. The quality factor $f_{quality}$ is formulated to maximize the number of low quality voxels visible from a given viewpoint.

Viewpoint space is taken to be a constant radius tessellated sphere, obtained by recursive sub-division of an icosahedron, fitted over the volumetric representation for the object. Using an experimental configuration whose positioning system was limited to a single degree of freedom, the authors show separate and combined plots of the objective function as a function of angle and the cumulative number of views taken, a presentation format usefully illustrating the impact of the quality measure.

While the introduction of a quality term in the objective function is a valuable contribution, the chosen measure is subjective and does not relate to objective measurement constraints on the reconstructed model. Further, the model is deterministic and considers only one error mechanism amongst several. Viewpoint space is constrained to two dimensions while experiments reported use a single (rotational) degree of freedom.

Octree Methods

Octree-based methods provide more efficient encoding of voxel occupancy.

Connolly One of the earliest papers on view planning is due to Connolly [Con85]. He appears to have first coined the term "next-best-view" (NBV). Connolly presents two algorithms, planetarium and normal, differing in the cue used to suggest feasible views and the selection mechanism used to find the next-best-view from the set of feasible candidates. The imaging workspace is voxelized and labelled as empty, occupied or unseen. This information is encoded in an octree which is also used to represent the object surface. All viewing vectors are assumed to point to the origin. Viewpoints are constrained to evenly spaced points on a sphere around the object.

The planetarium algorithm applies a visibility analysis of the octree for each candidate viewpoint on the sphere. The area of unseen voxels projected onto the image plane for each candidate viewpoint is taken as a measure of the solid angle of unseen space that will be swept by that view. The viewpoint with the largest unseen area is selected as the NBV. The algorithm suffers from time complexity problems inherent with a complete visibility analysis for all candidate viewpoints. The normal algorithm simplifies the visibility analysis to the local rather than global level by examining faces in the octree common to both unseen and empty voxels. It is faster but does not deal as well with self-occluding scenes.

The method is subject to many simplifications and limitations. Neither sensor nor positioning system performance is characterized. The sensor is treated as an ideal point source without constraints on field-of-view or depth-of-field. Shadow effects are ignored. Notwithstanding these limitations, Connolly's pioneering concept of exploring and labeling imaging workspace is found in many later papers in the field.

Space Carving

The space carving technique[4] applies primarily to a small class of range sensors with a limited sensing volume, particularly those designed as non-contact replacements for CMM mechanical touch probes. The sensor is swept through the imaging workspace in a pre-planned methodical manner, diverting around obstacles, with the objective of reliably labeling workspace occupancy. As with contour following, collision avoidance is a primary concern.

Papadopoulos-Orfanos A recent paper by Papadopoulos-Orfanos and Schmitt [POS97] summarizing Dimitri Papadopoulos-Orfanos' Ph.D. research [PO97] applies space carving to a specific shallow depth of field sensor for automated object reconstruction. Emphasis is placed on updating the voxel-based scene representation - in particular, robust labeling of empty space to avoid collisions. Voxel labeling is achieved by a complex analysis of occlusion zones followed by ray tracing operations. Most of the work is concentrated on collision avoidance and path planning.

Papadopoulos-Orfanos describes but does not fully implement a two stage 3D imaging strategy of scene exploration followed by surface data acquisition. The exploration stage, which is implemented, employs an exhaustive search of unknown portions of the workspace. In the vicinity of the object,

[4] As used here, "space carving" is distinct from the shape from silhouette technique [KS00].

the goal is to get the sensor as close as possible to the surface while avoiding collisions. During this phase, sensor orientation is fixed and only translations are employed to search the workspace layer by layer in a zigzag pattern. As a consequence of the fixed orientation, occluded surfaces are not imaged and surface occupancy in these regions is poorly defined. In a second stage of surface data acquisition (described but not implemented), sensor orientation view planning is proposed. Some existing NBV techniques are briefly discussed as candidate methods.

In its current form, the work does not address view planning per se. The exploration stage involves an exhaustive search of the imaging workspace following a pre-planned trajectory modified by collision avoidance. A second detailed imaging stage is briefly discussed but not implemented. No new view planning technique is proposed. If more fully developed, space carving has potential for high precision scanning. However, it will likely remain slow as a consequence of the small sensor frustum and the exhaustive search technique.

Solid Geometry Methods

This method utilizes standard solid geometry algorithms available with most CAD packages to model the current state of object knowledge. The method can be robust with respect to complex topology. A generic problem with the technique arises from solid geometry intersection operations which, by definition, subtract and cannot add volume. Therefore, if a range image used to extrude a solid volume does not completely cover the object, the resulting volume will exclude a portion of the object which can never be recovered by subsequent intersection operations. Consequently, difficulties arise when a view fails to completely enclose the object or along occlusion boundaries where data is often missing or erroneous due to inclination and edge effects.

Bi-Static Shadow Effects Although typical of depictions used in the literature, the previous illustration of shadow effects (Figure 2.3) is somewhat deceptive for two reasons. Firstly, it illustrates a mono-static sensor, whereas we are concerned with bi-static sensors whose optical baseline is generally significant with respect to the measurement stand-off distance. Therefore, the shadow effect is non-negligible and has two components - source and receiver shadows. Bi-static shadow effects are illustrated in Figure 2.4. We can observe that an occlusion edge as seen by the receiver may lie inside the object. This poses difficulties for occlusion-edge-based view planning techniques employing solid geometry intersection operations. Collision avoidance routines operating in conjunction with view planning also need to make allowance for this fact. Secondly, in general, the object will not lie entirely within the camera frustum so a given range image may or may not contain boundary occlusion edges. Some NBV algorithms can fail in the absence of occlusion edges.

Reed In a series of articles ([RA99], [RAS97a], [RAS97b], [RA97]) arising from his Ph.D. thesis, [Ree98], Reed uses a solid geometry approach to synthesize a continuous 3D viewing volume for a selected occlusion edge, rather than discretizing viewpoint space. A mesh surface is created for each range image and each mesh triangle is swept to form a solid model. Solid geometry union operations on the swept volume result in a composite model comprised of the surface and occluded volume, as known at that stage of modeling. Surfaces on the composite model are labeled "imaged surface" or "occluded surface".

Reed's view planning approach then proceeds with manual selection of a specific occluded edge as a target. A "visibility volume" is calculated for the target - that is, a 3D volume specifying the set of all sensor positions having

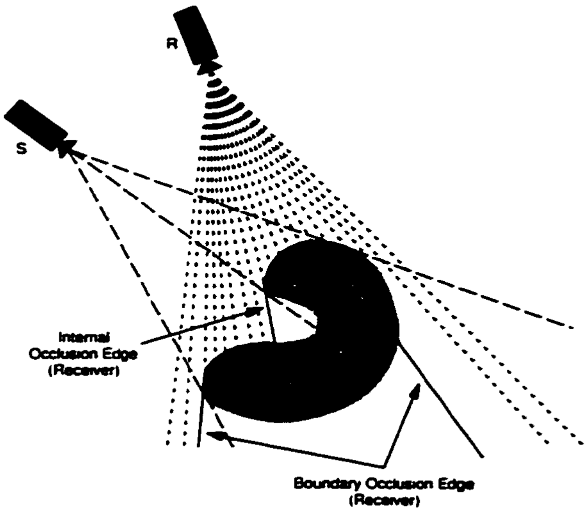


Figure 2.4: Occluding Edges in Geometric Images with Bi-static Sensor

an unoccluded view of the entire target for the model as presently defined. This is computed by subtracting the occlusion volume for each model surface component from the target's unoccluded volume, a half-space whose defining plane is coincident with the target's face and oriented in the appropriate direction. The work is based on previous sensor planning work in the same lab by Tarabanis et al [TTK96]. The volumetric approach allows sensor and positioning system constraints to be added.

The 3D imaging workspace search is an improvement over methods constrained to an arbitrary 2D surface around the object. The authors claim their approach has the advantage of defining more accurate viewpoints due to the treatment of viewpoint space as continuous rather than discrete, although no analytical or experimental data are provided to quantify this assertion.

Part of the reported research, [RAS97a], relies on manual selection of a suitable target and hence the NBV. There is no discussion of the sensitivity of the algorithm to target occlusion edge selection. A second paper [RAS97b] discusses an approach to automated viewpoint selection. In this case, the algorithm selects the NBV as the one imaging the most occluded surface elements in the current composite model. The NBV is determined by computing the visibility volume for each occluded surface element, intersecting each visibility volume with the sensor's reachable space and searching the intersection for the point imaging the most surface area. A third paper [RA99] briefly examines NBV planning based on the use of multiple targets.

Benefits of solid geometry methods include robustness to complex object topology and the guarantee of water-tight models. Viewpoint synthesis can be computationally advantageous with respect to methods which discretize viewpoint space. However, as presently defined, the approach has limitations. The sensor frustum is only partially modeled - field of view constraints are not considered. Bi-static shadow effects on view planning are not addressed directly; although missing data in small shadow zones is dealt with by interpolation and along occlusion boundaries by surface extension along the direction of the sensor baseline. Sensor errors are only briefly considered. View overlap planning is not addressed. Set operations on extruded surface elements introduce artifacts along integration boundaries.

2.3.3 Global View Planning Methods

A few methods derive a view planning cue from global rather than local characteristics of the geometric data.

Mass Vector Chain

In [Yua95], [Yua93], Yuan describes an interesting view planning mechanism. He observes that "a reconstruction system expects a self-controlled modeling

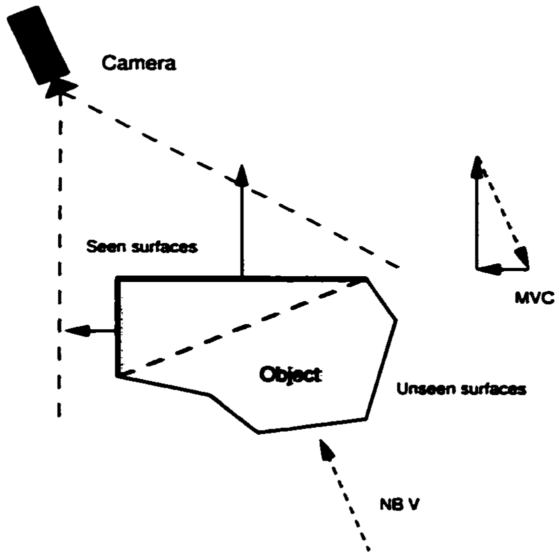


Figure 2.5: Mass Vector Chain

mechanism for the system to check the spatial closure of object models and to estimate the direction of unprocessed features".

Illustrated in Figure 2.5, his approach segments the observed surface into a set of small patches and describes the object with a mass vector chain (MVC). By definition, a MVC is a series of weighted vectors. A mass vector $\vec{V_i}$ is assigned to each surface patch S_i of the object. Thus, $\vec{V_i} = \vec{n_i}R_i$, where $\vec{n_i}$ is the average visible direction and R_i is the surface size when viewed from that direction. It is easily shown that the boundary surfaces of an object compose a closed surface boundary only when their mass vectors form a closed chain. Thus, the next-best-view can be set to the negative of the cumulative MVC, which should define a viewing direction whose mass vector will close the chain or at least shorten the gap. Special rules are required for holes and cavities.

The algorithm is shown to work with synthetic geometric data for a simple object. However, the idealized imaging environment does not address the complications of modeling automation with real sensors and complex man-made or natural objects. The method is capable only of estimating viewing direction, not position. It has no reliable measure of scale, range or the imaging volume. The method is able to deal with moderately complex topology and may have potential for machine vision tasks involving man-made parts with simple shape or for the initial global view planning phase of a two step coarse-fine planning process.

Intermediate Space Representations

The essence of view planning is capturing, representing and optimizing measures of visibility of the object surface from sensor poses in viewpoint space. This mapping between the object surface and points in the workspace of the sensor and positioning system involves a large amount of information and therefore a high degree of computational complexity for its acquisition, storage and manipulation. It is desirable to find a more compact and easily searched representation of the visibility information with a minimum loss of information. One approach following this line of reasoning involves encoding visibility information on a virtual surface positioned between the object and the sensor workspace - that is, an *intermediate space representation*.

Positional Space - Pito Richard Pito's 1997 thesis [Pit97a] is significant for its contributions to all four aspects of geometric modeling automation (scanning, registration, integration and view planning). See also the summary paper [Pit99] as well as [Pit97b], [Pit96b], [Pit96a], [PB95].

Pito's work is unusual in that he first characterizes performance of the range scanner used for experiments and then incorporates sensor models in the automation algorithms. In addition to treating the grazing angle effect (fairly common in the literature), he also treats edge effect anomalies during the integration phase. The view overlap requirement is incorporated. The shadow effect is explicitly treated. However, non-stationary measurement error within the measurement volume is not addressed. The need for generalized viewpoints is discussed but apparently not implemented. The technique is amenable to using generalized viewpoints, at the expense of increasing the dimensionality of intermediate space.

Like many others before him, Pito chooses occlusion edges as a cuing mechanism for the NBV search. He observes that the void volume can be economically represented by defining only the void surface near edges of the current model, which he represents by small rectangular patches attached to occluding edges of the seen surface. He argues that it is unnecessary to represent the complete boundary of the umbra cast by the object. In general, the portion of the umbra nearest the occlusion boundary will be nearest to the real object surface and therefore be the best region to search next. Furthermore, such regions fit the overlap constraint.

As illustrated in Figure 2.6, Pito's NBV algorithm is based on using an intermediate space representation, "positional space" (PS), as a repository for two types of visibility information - object surface visibility and the sensor's scanning potential. A virtual positional space surface (PSS), with a shape appropriate to the sensor-positioning system combination, is placed between the object and the sensor workspace. Object surface space (represented as a mesh), PSS and viewpoint space are discretized.

The visibility of each triangular mesh element on the object surface can be encoded in positional space by tracing an "observation ray" from the mesh element to a PSS cell. The direction of an unoccluded observation ray relative to a local frame of reference on the PSS can then be specified

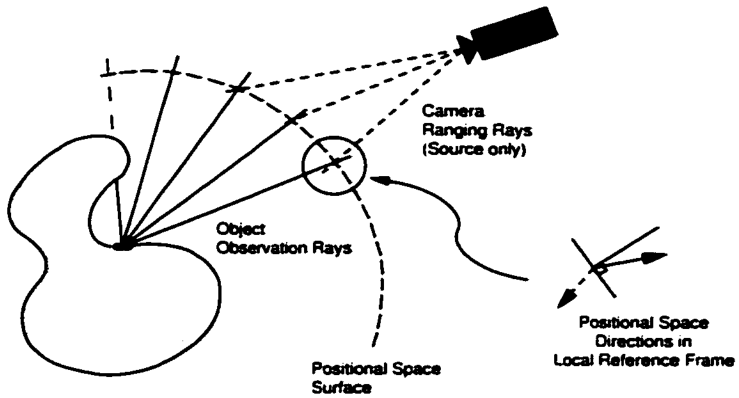


Figure 2.6: Positional Space

by two angles, termed the positional space direction (PSD). Pito chooses to encode surface visibility as an analogue value equal to the area of the surface element visible by a given ray weighted by a confidence measure tied to the measurement grazing angle. Thus, in Pito's concept, PS is a scalar field in four dimensions $P(u, v, \theta, \phi)$. Encoding the image of both the seen surface and void patches in PS provides a means to apply an overlap constraint between views to meet registration and integration requirements. The range camera's scanning potential at a given viewpoint can be similarly encoded in positional space by determining the intersection of "ranging rays" from the optical transmitter and receiver with the PSS. A separate image must be calculated for each viewpoint. Using PS as a placeholder for ranging and observation rays facilitates determining which or how many of them are collinear, as well as aiding the application of NBV constraints.

Without detracting from the comprehensive treatment and contribution of the work to the field, there are some unresolved issues and weaknesses with the approach:

- Overall computational complexity issues are not addressed; nor are trade-offs between discretization levels in object surface space S, viewpoint space V or positional space PSS. In the end, visibility analysis relates S to V. The impact of quantization errors arising from imposition of an intermediate space between S and V is not addressed. Efficiency is a concern. The experimental setup used to demonstrate the algorithm was limited to a single rotational degree of freedom.

- One of the major claims of the work is that the computational burden of visibility analysis is decoupled from the size of viewpoint space. Nevertheless, ray tracing figures heavily in the approach and it has not been demonstrated that the net visibility analysis burden is lowered for a given performance level. However, the algorithm appears to be readily parallelizable.

- PS representation is a lossy compression of pose information in which range-related information is lost. Stand-off range critically impacts sensor performance.

- A PS image must be cast for each viewpoint, a potentially large number. The algorithm does not provide guidance on efficiently sampling viewpoint space. On the other hand, PS sensor images need be calculated only once and stored off-line. Memory issues are not addressed.

- Reference is made to the use of multiple PS surfaces which would capture additional information in viewpoint space but this would appear to reduce the claimed computational advantages of the representation.

Expert System

Artificial intelligence and expert system approaches to view planning have been few and limited mainly to illumination issues [Bat89], [Nov88], [KST90].

2.4 Comparison of Existing Methods

Recall that:

- Maver and Bajscy select a NBV by reasoning about the angular arc through which occluded areas of the object surface can be viewed, based on assumptions about the unseen topography.

- Connolly 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.

- Yuan selects a next best viewing direction based on a global analysis of the mass vector chain for the object surface.

- Banta et al discuss several view planning mechanisms, the most recent being the NBV oriented to the centroid of the cluster containing the largest number of unknown voxel faces.

- Massios and Fisher use an objective function incorporating both visibility and "quality" terms.

- Tarbox and Gottschlich compute a measurability matrix encoding a complete visibility analysis over the set of all viewpoints and surface elements, and select a viewpoint maximizing the void surface visible.

- Reed et al first synthesize a visibility volume for a selected occluded edge using solid geometry techniques and then select a NBV imaging the most occluded surface elements.

- Whaite and Ferrie use uncertainty in fitting parameterized superquadrics to segmented range data as a cue to viewpoint selection and then select a viewpoint at a fixed offset in the direction in which the variance of the fit of the data to the current model is the greatest.

- Pito separates the visibility analysis of the object surface from that for the sensor by casting visibility images onto an intermediate space representation, Positional Space, and optimizes viewing direction selection within that domain subject to various constraints.

- Papadopoulos-Orfanos defines a space carving technique for a shallow depth of field range sensor based on a pre-planned search pattern plus collision avoidance.

Table 2.2 summarizes attributes of existing approaches to view planning with respect to the view planning task requirements specified in Chapter 1. The legend for the comparison table is as follows: requirement satisfied (Y), not satisfied (N), partially satisfied (P), uncertain (?) or not applicable (−). The research objective is a column of all Y's. In chapter 8, we evaluate the view planning approach presented in this thesis according to the same criteria.

Category	Requirement	Maver & Bajscy	Connolly	Yuan	Banta et al	Massios & Fisher	Tarbox & Gottschlich	Reed et al	Whaite & Ferrie	Pito	Papadopoulos-Orfanos
General	Model quality specification	N	N	N	N	N	N	N	N	N	N
	Generalizable algorithm	N	Y	N	Y	Y	Y	Y	N	Y	N
	Generalized viewpoints	N	N	N	N	N	N	N	N	N	N
	View overlap	N	N	N	N	N	N	N	P	P	N
	Robust	N	N	N	N	?	?	?	N	?	?
	Efficient	N	?	?	?	?	N	?	N	?	N
	Self-terminating	N	Y	Y	Y	Y	Y	Y	Y	Y	Y
Object	Minimal a priori knowledge	Y	Y	Y	Y	Y	-	Y	N	Y	Y
	Shape constraints	N	Y	N	Y	Y	Y	Y	N	Y	Y
	Material constraints	N	N	N	N	N	N	N	N	N	N
Sensor	Frustum modeled	N	N	N	N	N	N	P	N	P	Y
	Shadow effect	Y	N	N	N	N	Y	N	N	Y	Y
	Measurement performance	N	N	N	N	P	P	N	P	P	N
Positioning System	6D pose	N	N	N	N	N	N	N	N	N	Y
	Pose constraints	N	N	N	N	Y	Y	Y	N	Y	Y
	Positioning performance	N	N	N	N	N	P	N	N	N	N

Table 2.2: Comparison of View Planning Algorithms

2.5 Critique of Existing Methods

2.5.1 General

Model quality specification None of the methods reviewed incorporate explicit quality goals for the reconstructed object model. Massios incorporates a quality term in the view planning objective function but this term is not tied to explicit pass/fail criteria. Several authors include a "measurement quality" term tied to grazing angle but do not relate the factor to explicit model performance requirements.

Generalizable algorithm Several methods are generalizable to a broad class of range sensors, positioning systems and objects. Others are not.

Generalized viewpoint Several authors describe a generalized viewpoint in their opening theoretical problem formulation. However, none retain the formulation in the actual algorithm development. In most cases, pose space is constrained to one or two dimensions versus the generalized case which is $6D^+$. No author considers a programmable sensor parameter other than pose. Some methods (Yuan, Pito) are capable of providing only incomplete viewpoint information, specifically viewing direction but not sensor position or orientation about the sensor boresight.

View overlap Almost all authors assume an error-free positioning system and ignore view planning constraints for view overlap for image-based registration. Only Pito applies a view overlap constraint, although shape complexity is not addressed.

Robust None of the methods incorporate both sensor and positioning system error models. Several methods (Massios, Tarbox, Pito) consider grazing angle effects as a subjective quality factor. No method considers geometric step edges, reflectance step edges or multiple reflection effects. The robustness of the proposed methods in realistic sensing and view planning environments is either weak or is undetermined - that is, the environment of real sensors with real errors and artifacts, real objects with representative topological and geometrical complexity, and real positioning systems subject to pose error and restricted movement envelopes.

Efficiency Given the lack of standardized performance benchmarks, it is difficult to make quantitative performance comparisons between proposed view planning methods. However, some authors (Banta, Tarbox, Connolly) describe multiple algorithms and provide qualitative performance comparisons. Few authors explicitly evaluate the computational complexity of their algorithms. Computational complexity is an issue with all methods. However, many of the techniques are parallelizable and computational power and memory are becoming less of a concern with ongoing advances in computer technology.

Self terminating Various termination criteria are used, but none relate to explicit requirements to be met by the reconstructed object model.

2.5.2 Object Characterization

Minimal a priori knowledge Most methods impose no a priori knowledge constraints on the objects to be modeled. Tarbox's work focussed on the inspection application. Whaite and Ferrie's approach was aimed at robotic environment exploration.

Shape constraints Similarly, most methods impose no shape constraints on the objects to be modeled. However, some would clearly be unable to handle complex shapes. Performance was not characterized with respect to increasing object topological or geometrical complexity.

Material constraints No method dealt with object material effects on sensor performance.

2.5.3 Sensor Characterization

Frustum modeled Most methods do not incorporate a realistic model of the sensor frustum i.e. the depth-of-field, angular field-of-view and scan length/arc of real range cameras.

Shadow effect Few methods model shadow effects. This core attribute of triangulation-based range sensors limits observation in cavities or around obstacles. A great deal of manual view planning effort is spent addressing shadow zones.

Measurement performance characterization With some exceptions ([Pit99]), the view planning literature generally assumes idealized, error-free sensors. Many model the camera as an ideal mono-static sensor. Yet, geometric measurement techniques are highly non-linear and error prone. A substantial portion of the effort in manual object reconstruction is spent minimizing errors in data acquisition and model building. The sensing error most commonly considered in the literature is inclination bias. No author addresses the non-isotropic and non-stationary nature of residual geometric noise within the sensor frustum of a calibrated range sensor. Traditional view planning approaches have attempted to obtain complete coverage of the object surface with limited attention to the quality of the geometric data.

2.5.4 Positioning System Characterization

6D pose Most methods limit viewpoint space to one or two dimensions. The most common configuration constrains the sensor to a "viewing sphere" with the camera boresight fixed on an object-centered origin. It is not possible to optimize sensor performance with such limitations.

Pose constraints Several of the more advanced approaches to view planning incorporate positioning system constraints.

Positioning performance Excepting Tarbox, no author addresses the impact of pose error on view planning. Tarbox removes viewpoints not robust to pose variation from the viewpoint set associated with each surface point. Positioning system performance is not modeled.

2.5.5 Critique Summary

Techniques and Representations By far the most common NBV cuing mechanism is either the surface or the solid volume of the umbra cast by a range image view of the scene - that is, occlusion edges or the shadow zone. The greatest variability between view planning approaches lies in the mechanism used to select the NBV. Common surface representations are volumetric occupancy and surface meshes. The most common representation of imaging workspace is volumetric - voxel occupancy, octree or solid geometry.

Object Size Assumption Many techniques assume that the object falls completely within the camera field of view and would fail if this condition is not met. For example, if a range image contains no jump edges, what should the algorithm do next?

Strategies Most approaches were monolithic in the sense that a single technique was used throughout. It would appear advantageous to divide the overall view planning problem into stages as Hall-Holt suggests and consider different techniques at different stages - in particular, restricting the use of computationally intensive techniques to limited subsets of the overall problem. All of the methods examined are deterministic. There may be opportunities for random selection methodologies.

Performance Objectives and Evaluation None of the current methods include explicit pass/fail quality criteria on the reconstructed model. Generally, there is a lack of clarity with respect to the view planning objective. Is view planning a search for some global optimality or is it a search for merely acceptable viewpoints? What determines success - a minimum number of viewpoints; time to complete the job; some other criteria?

Suitability As presently developed, none of the traditional view planning methods is directly suitable for performance-oriented automated object reconstruction for the following principle reasons: overly constrained viewpoint space, inadequate sensor and positioning system performance characterization and excessive computational complexity.

2.6 Open Problems

The open problems in view planning for automated object reconstruction are intriguingly simple and yet exceptionally complex at the same time. There is as yet no general purpose solution. While a variety of techniques have been described in the open literature, none fully meets the requirements for performance-oriented view planning. It is notable that there are no known automated geometric modeling environments or performance-oriented view planning systems in general public or commercial usage.

There are three main open issues with respect to performance-oriented view planning for automated object reconstruction. These are efficiency, accuracy and robustness.

2.6.1 Efficiency

The efficiency issue relates to the computational complexity of the view planning algorithm in terms of both time and memory, although timeliness is probably the more important factor. No method proposed to date is able to provide both performance and efficiency. In fact, even with unrealistic constraints on the problem domain, most view planning techniques are unacceptably slow. However, ongoing computer technology developments are making this less of a concern.

2.6.2 Accuracy and Robustness

Accuracy and robustness considerations go hand-in-hand. Both demand that the view planning module incorporate reasonably complete sensor noise and artifact models based on performance characterization of the sensor and positioning system. Additionally, high accuracy object reconstruction requires a highly mobile positioning system. Desirably, this means a full six degrees of freedom over an adequate imaging workspace. View planning algorithms and imaging environments restricted to a one or two dimensional viewpoint space can at best provide limited-quality, limited-coverage models for a restricted class of objects.

2.6.3 Other Considerations

The technology development experience to date suggests we adopt a more flexible, multi-faceted approach to the problem and consider:

(1) a multi-phase approach such as scene exploration, rough modeling, fine modeling, problem resolution,
(2) use of multiple techniques in combination instead of a monolithic approach using a single technique,
(3) heuristic and expert system approaches emulating human operators, and
(4) multi-sensor fusion approaches - perhaps using a fast, wide field-of-view, low precision sensor for scene exploration and collision avoidance in combination with a high quality range sensor for precise surface measurements.

Additionally, it is evident that the field presently lacks the means for meaningful comparative performance assessment and could benefit from standardized performance benchmarks and quantitative algorithm performance analysis. Finally, to date there has been only limited treatment of the theoretical aspects of the view planning problem.

Chapter 3

Modified Measurability Matrix

This chapter introduces the modified measurability matrix (3M) algorithm for model-based performance-oriented view planning. The concept is refined and more fully examined in subsequent chapters.

3.1 Model-Based View Planning

Object reconstruction requires two functions to be performed - scene exploration and precision measurement. The first, scene exploration, is necessary because we begin the problem with limited information. Typically, this includes the approximate bounding dimensions and center of the object and perhaps its overall topology. The second, precision measurement, is directly related to reconstruction quality. Traditional view planning methods (Figure 3.1) attempt to undertake these functions simultaneously. They take an image, acquire some new information, augment the partial scene model, decide on the NBV and repeat the process. Such methods are intrinsically incremental. The associated irrevocable imaging strategy is inherently sub-optimal in terms of the number of views required.

Multi-stage model-based view planning (Figures 3.2, 3.3) takes a differ-

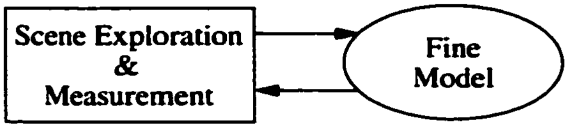


Figure 3.1: Traditional View Planning

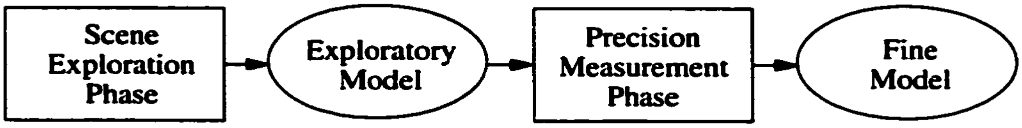


Figure 3.2: Model-Based View Planning

ent approach by separating scene exploration from precision measurement. Based on limited knowledge contained in an *a priori model*, the process begins with a rapid, preprogrammed scene exploration phase to acquire an *exploratory or rough model*. As long as the exploratory model captures the essence of the object's shape, it can have relatively low sampling resolution and accuracy. The exploratory model is then used to plan precise, dense surface scanning of the object to produce the desired high quality reconstruction, the *fine model*. With the benefit of the overall scene knowledge embedded in the rough model, the precision measurement phase has the option of several NBV selection strategies. If required by the application, this includes the ability to use a more efficient revocable selection strategy to produce an optimal or near-optimal view plan. The rough model may be discarded once it has served its purpose. In our approach, we use polygonal mesh model representations ([HB94]) for both the rough and fine models to take advantage of their inherent flexibility and ease of manipulation.

Although additional stages could conceivably be utilized, the current con-

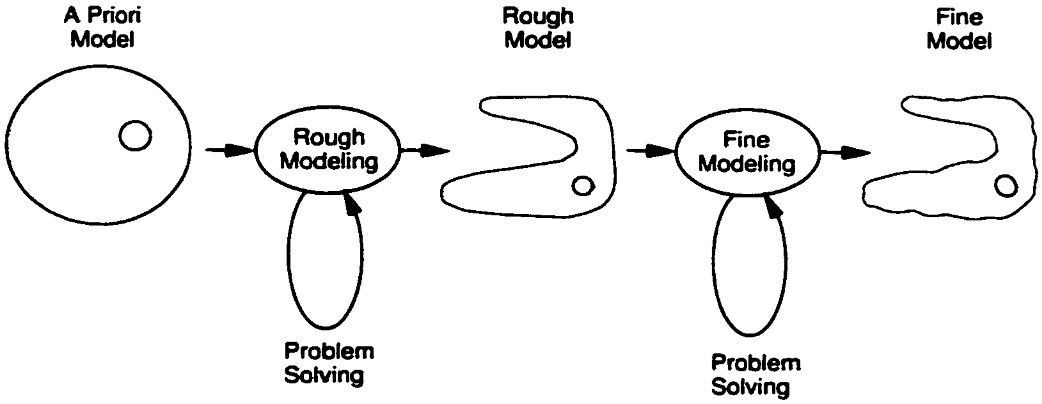


Figure 3.3: Multi-stage Model-based View Planning

cept (Figure 3.3) involves a two step object reconstruction process, separating the exploration and precision measurement processes. Problem solving subfunctions address model quality discrepancies at each stage. This multi-stage model-based view planning strategy is one 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 meeting 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 specified requirements is achieved. The approach is also incremental, but at the model rather than the image level. Within each level, the process is inherently parallel.

View planning for inspection applications, such as [TG95], [PRBL99] is intrinsically model-based as it starts with a CAD model of the object to be inspected. A few approaches to automated or semi-automated object reconstruction utilize an initial low resolution full or partial object scan for

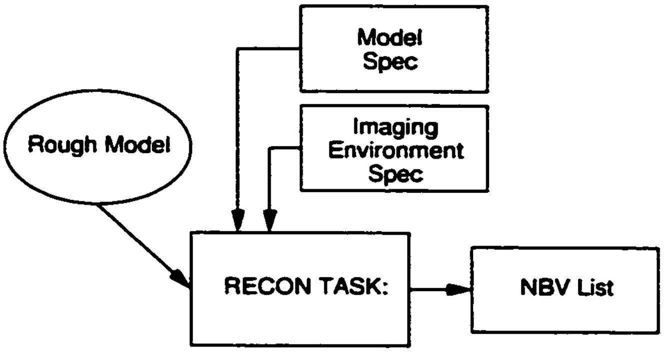


Figure 3.4: Performance-Oriented View Planning

subsequent view planning purposes. Variations of contour following schemes used by Soucy [SCF98] and Lamb [LBG99] follow this approach. Garcia [GVSB98], [GVS98] also uses a two stage algorithm in which the first stage attempts to sample most of the object using an occlusion edge method. Computationally expensive visibility analysis is left to a second hole-filling stage.

3.2 Performance-Oriented Reconstruction

Our approach is also model-based in a second sense. In the introduction, we defined *performance-oriented* reconstruction as model acquisition based on a set of explicit quality requirements expressed in a *model specification*. In addition to all-aspect coverage, measurement quality is specified in terms of precision, sampling density and perhaps other quality factors. The specified measurement quality may be fixed or variable over the object surface. Performance-oriented view planning requires suitable models of both sensor and positioning system performance (Figure 3.4). These can be combined in an imaging environment specification. Specifically, it requires:

- a *sensor model* to include a description of the camera geometry, frustum geometry and characterization of measurement performance within the calibrated region, and
- a *positioning system model* describing the degrees of freedom, range of motion and positioning performance within the movement envelope.

There has been relatively little prior work on performance-oriented view planning. Cowan and Kovesi [CK88] use a constraint satisfaction approach for sensor location subject to task requirements which include resolution, focus, field of view, visibility, view angle and prohibited regions. The MVP system developed by Tarabanis et al [TTA95] automatically synthesizes views for robotic vision tasks with intensity cameras based on a task specification as well as a geometric model of the scene and optical models of the sensor and illumination sources. The MVP task specification includes visibility, field of view, resolution, focus and image contrast. Soucy et al [SCF98] describe a system for automatically digitizing the surface of an object to a prescribed sampling density using a contour following scheme. Prieto [PRBL01], [PRBL99] sets CAD-based inspection criteria based on range sensor performance characterization. Additionally, as noted in Chapter 2, a number of authors have considered grazing angle as a subjective quality measure. Otherwise, the objective of most traditional work in the view planning field has been explicitly or implicitly limited to full surface coverage.

3.3 Measurability Matrix

3.3.1 Data Structure

Before proceeding further, it is pertinent to examine the basic data structure used in model-based view planning. A measurability matrix $\mathbf{M}$ (Figure 3.5) is

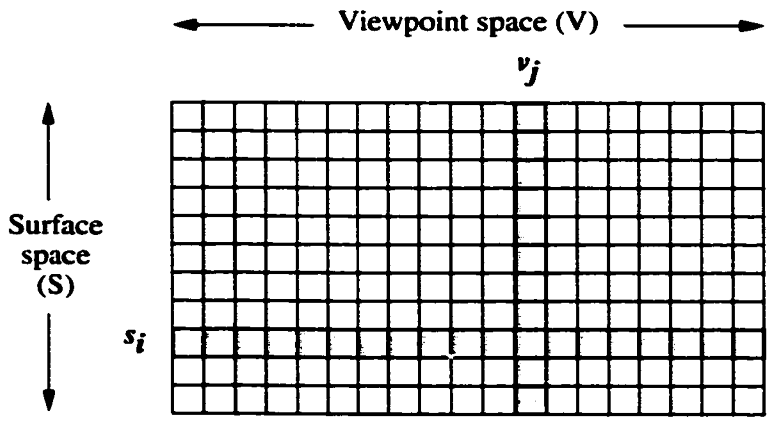


Figure 3.5: Measurability Matrix

a two dimensional array of measurability elements m_{ij}. By convention, rows correspond to surface points and columns to viewpoints. Thus, rows of $\mathbf{M}$ span discretized object surface space S (vertices of the rough model mesh in our case) while the columns span discretized viewpoint space V. Both binary and analogue measurability matrices can be defined. The binary version is more compact and best suits the purposes of the current work. Each element m_{ij} of $\mathbf{M}$ is a binary estimate of the measurability of a single surface point $\mathbf{s}_i$ on the exploratory model from a single viewpoint $\mathbf{v}_j$. The estimate serves to predict scanning results at a higher sampling density with the real sensor, real object and real imaging environment.

As originally formulated by Tarbox [TG95], each measurability element was required to pass visibility tests for both optical source and receiver as well as thresholds on grazing angle. In our case, we apply more stringent requirements related directly to the model specification for the reconstruction task. These tests explicitly incorporate the task specification as well as performance models for the range camera and positioning system. To be declared measurable, a boolean measurability matrix element m_{ij} must pass the following tests:

- *Frustum occupancy* - Surface point $\mathbf{s}_i$ must lie within the world space sensor frustum as defined by viewpoint $\mathbf{v}_j$.

- *Visibility* - Surface point $\mathbf{s}_i$ must be visible to both the world space optical source and receiver positions as defined by viewpoint $\mathbf{v}_j$.

- *Specification-compliance* - The estimated measurement precision and sampling density for surface point $\mathbf{s}_i$ as sampled by viewpoint $\mathbf{v}_j$ must meet the requirements of the model specification.

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. 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. These relationships are illustrated in Figure 3.6 with 2D slices through an object resembling a coffee cup.

From the foregoing, we can observe that the solution to the view planning problem is simply to cover the rows of $\mathbf{M}$ with a minimal subset of its columns. Consequently, it is immediately apparent why the view planning problem is isomorphic to the set covering problem (SCP). The SCP is a well-studied problem in combinatorial optimization that is known to be NP-complete [GJ79].

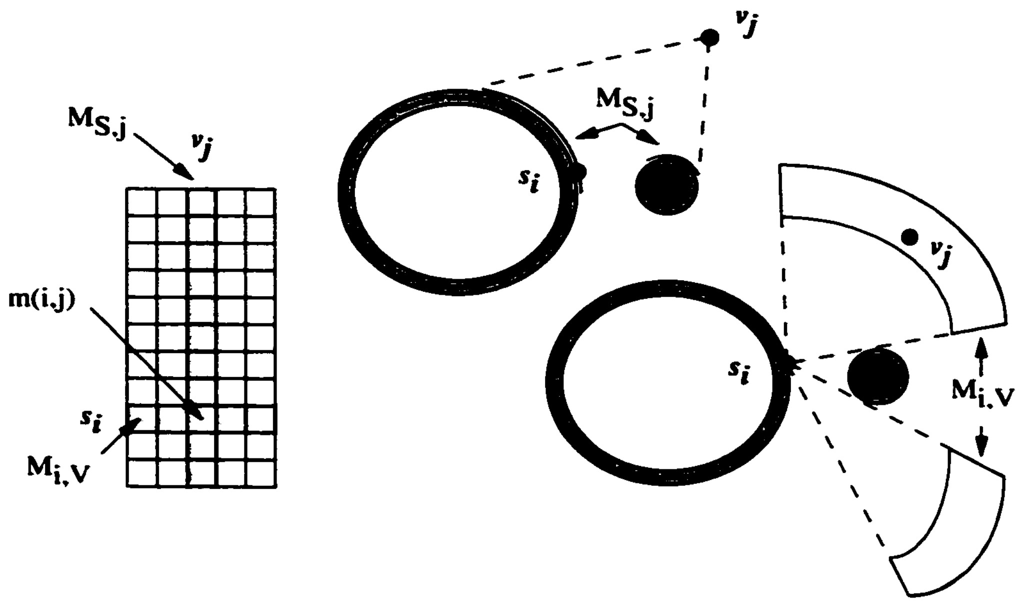


Figure 3.6: Measurability Matrix Row and Column Vectors

3.3.2 Measurability

We use several variants of the term "measurability". Each *measurability matrix element* m_{ij} is a binary estimate of the measurability of a single surface point $\mathbf{s}_i$ on the rough model from a single viewpoint $\mathbf{v}_j$. The estimate serves to predict scanning results at a higher sampling density with the real sensor, real object and real imaging environment. To be measurable, all model specification requirements must be met at that surface point for that viewpoint.

The *measurability of a single viewpoint* $m(\mathbf{v}_j)$ or, equivalently an image, is the ratio of the surface coverage of that viewpoint $|S_j|$ relative to the size of the surface patch as a whole $|S|$. 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, that is

$$m(\mathbf{v}_j) = \frac{|S_j|}{|S|}. \tag{3.1}$$

Similarly, the *measurability of a set of viewpoints* is the ratio of the joint coverage of those viewpoints to the size of the surface patch as a whole -

$$m(\mathbf{v}_j, \mathbf{v}_k) = \frac{|S_j \cup S_k|}{|S|}. \tag{3.2}$$

A *measurability projection* is a 2D projection of a labelled rough or fine model mesh whose vertices are encoded with their binary measurability. For display purposes, these can be coloured white for measurable and black for non-measurable. Refer to Figure 4.19 for example measurability projections.

A *measurability image* is a range image in which the pixels are labelled with their binary measurability. For display purposes, these can be coloured white for measurable and black for non-measurable. A measurability image is the labeling of a set of range samples covered by a single frustum cross-section associated with a single viewpoint while a measurability projection is the labeling of an entire object model resulting from the coverage of several range images taken from different viewpoints.

The view planning algorithm operates on *estimated measurability* of a single viewpoint or sets of viewpoints. When the resulting NBV list is executed by a real sensor against the real object or by a sensor model against a simulated fine model, the result is *verified measurability*. Finally, the term *composite measurability* will be used with both estimated and verified results to define the net coverage of viewpoint sets.

3.4 Sensor Performance Model

In this section, we present a general purpose performance model for triangulation-based range cameras as used in the 3M algorithm. The model is subject to appropriate adjustments to fit the geometry and scanning mechanism of each specific class of range camera. The 3M algorithm also addresses pose errors resulting from positioning system imperfections but we delay presentation of a generic positioning system error model until Chapter 7.

The sensor performance model embedded in the 3M system is reasonably comprehensive. Several generic scanning mechanisms and their associated frustum shapes are defined and parameterized in terms of field of view, depth of view and scan length or arc. The camera is modeled as a bi-static sensor with a specified optical baseline. The origin of the camera co-ordinate reference frame is taken to be the location of the laser source at its nominal rest position. Image size and other parameters are optionally reconfigurable. Viewpoints are treated as generalized viewpoints $(\mathbf{v}, \lambda_s)$, consisting of sensor pose $\mathbf{v}$ and a set of controllable sensor parameters λ_s. Thus each viewpoint has a defined frustum in camera space and an associated sensor configuration.

As currently used, the models for measurement precision and sampling density incorporate effects related to bi-static visibility, standoff distance, scanning angle and grazing angle. In the future, the scope and sophistication of these sensor models and the associated measurability matrix compliance tests can be expanded to include other measurement error factors.

3.4.1 Range Camera Geometry

To illustrate the sensor performance model, we examine the case of a line-scan range camera, a common configuration whose imaging geometry is shown at Figure 3.7. Following the convention in the field, the camera axis defines the

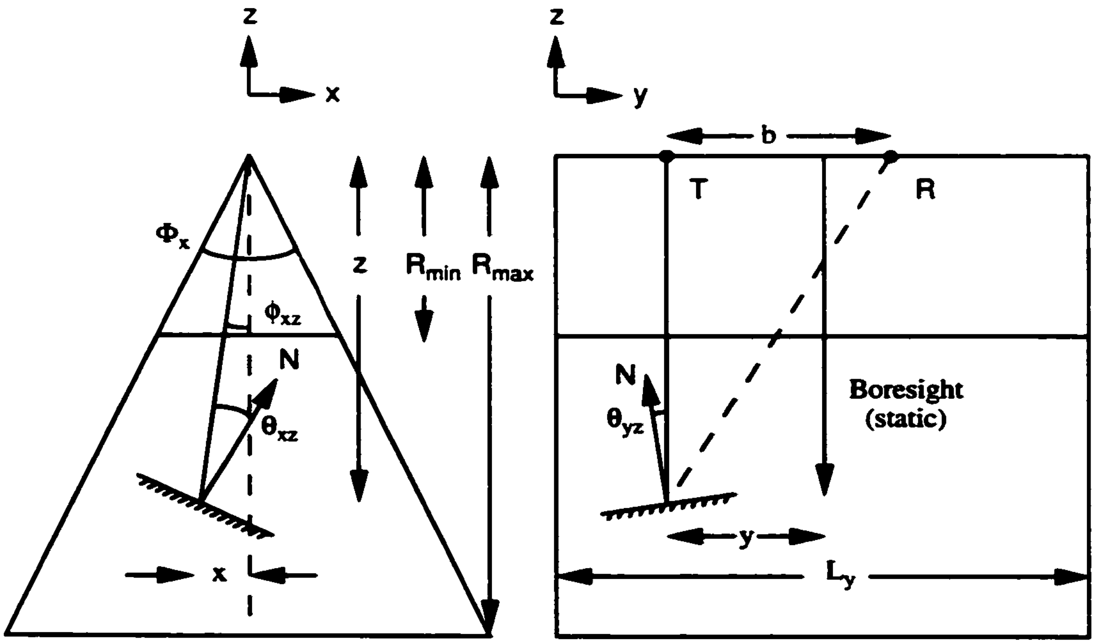


Figure 3.7: Line-scan Range Camera Geometry

negative z-axis. The negative sign is dropped when referring to range along the z-axis, provided the situation is clear. The frustum is defined by Φ_x (the sensor angular field of view in the x-z plane), L_y (the linear scan length in the y-z plane) and R_{min} and R_{max} (the minimum and maximum scanning ranges along the z-axis). The optical transmitter (laser) and optical receiver (detector) are separated by a distance equal to the optical baseline b along the y-axis. In the x-z plane, scanning occurs optically, while in the y-z plane it occurs through mechanical movement of the camera along the y-axis. ϕ_{xz} is the instantaneous laser scan angle in the x-z plane. In our convention, the origin of the sensor frame of reference is located half way along the linear scan and defines the nominal rest position of the laser.

3.4.2 Measurement Precision

Our sensor model incorporates a model of geometric noise based on range camera calibration data [BEC93], [EB94], [EB95]. The noise model estimates statistics of the residual random noise remaining after calibration. We use the standard convention for the range camera coordinate frame of reference: the x, y and negative z-axes are parallel to the laser scanning plane, camera optical baseline and camera boresight, respectively. Sensor geometric noise effects are modeled as follows, where $\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z$ are the standard deviation estimated noise components along the respective sensor axes, noise coefficients (C_x, C_y, C_z) are derived from calibration data and z is range along the camera boresight:

$$\begin{aligned} \hat{\sigma}_x &= C_x z, \\ \hat{\sigma}_y &= C_y z, \\ \hat{\sigma}_z &= C_z z^2. \end{aligned} \tag{3.3}$$

In this first order model, off-diagonal noise co-variance matrix elements are considered secondary effects ignored for view planning purposes.

As noise along the sensor boresight predominates and small errors tangential to the surface are of debatable significance, we use $\hat{\sigma}_z$ as a surrogate for measurement precision. Additionally, estimated sensor noise is modified by an experimentally-based grazing angle model for the sensor x-z and y-z planes. Incidence angle effects are most noticeable in the plane of triangulation - that is, the yz-plane, where they generally follow an inverse cosine relationship up to a cut-off angle t_{yz}. There is no noticeable inclination effect in the scanning plane up to a cut-off angle t_{xz} at which point the received energy drops below threshold. Typical grazing angle threshold values are

$t_{yz} = 60°$ and $t_{xz} = 70°$. Thus, we modify the estimated precision obtained from Equation 3.3 by the following, where $U_i(\theta)$ is the inverse unit step function, i.e. $U_i(\theta) = 1 - U(\theta)$:

$$\hat{\sigma}_z = \frac{C_z z^2}{\cos(\theta_{yz})\; U_i(\theta_{yz} - t_{yz})\; U_i(\theta_{xz} - t_{xz})}. \tag{3.4}$$

While grazing angle effects are frequently modeled in view planning algorithms ([BA96], [MF98], [TG95], [Pit99]), they have been treated subjectively rather than objectively. That is, low grazing angle viewpoints are preferred over high grazing angles, rather than evaluating the objective effects of grazing angle on measurement quality, such as precision and sampling density.

3.4.3 Sampling Density

The imaging geometry of a line-scan range camera was shown at Figure 3.7. We use a conservative chord-based estimate for sampling density $\hat{\rho}_z$ where δx and δy are the estimated sampling intervals along the sensor x- and y-axes and δc is the estimated inter-sample chord length. Then,

$$\hat{\rho}_z = \frac{1}{\delta x^2 + \delta y^2} = \frac{1}{\delta c^2} \tag{3.5}$$

where

$$\delta x = R_{xz} \frac{\Phi_x}{N_x - 1} \frac{1}{\cos \theta_{xz}} \tag{3.6}$$

and

$$\delta y = \frac{L_y}{N_y - 1} \frac{1}{\cos \theta_{yz}}. \tag{3.7}$$

In Equation 3.6, $R_{xz} = z/\cos\phi_{xz}$ is the slant range, $\Phi_x/(N_x - 1)$ is the angular sampling interval in the x-z plane and $1/\cos\theta_{xz}$ is the inclination effect in the x-z plane. In Equation 3.7, $L_y/(N_y - 1)$ is the linear sampling interval in the y-z plane and $1/\cos\theta_{yz}$ is the inclination effect in the y-z plane. Image size is N_x-by-N_y samples. For simplicity, we adopt the abbreviated notation $S\theta = \sin\theta$, etc. Then, the estimated sampling density is

$$\hat{\rho}_z = \frac{(N_x - 1)^2(N_y - 1)^2C^2\theta_{xz}C^2\theta_{yz}}{{R_{xz}}^2{\Phi_x}^2(N_y - 1)^2C^2\theta_{yz} + {L_y}^2(N_x - 1)^2C^2\theta_{xz}}. \tag{3.8}$$

In many cases, Equation 3.8 can be simplified for symmetric range images by setting $N = N_x = N_y$, giving

$$\hat{\rho}_z = \frac{(N - 1)^2C^2\theta_{xz}C^2\theta_{yz}}{{R_{xz}}^2{\Phi_x}^2C^2\theta_{yz} + {L_y}^2C^2\theta_{xz}}. \tag{3.9}$$

3.5 3M Algorithm

3.5.1 Concept

Overview We have seen from the literature survey of the preceding chapter that the main open issues are accuracy, robustness and efficiency. The modified measurability matrix (3M) view planning algorithm attempts to address these open issues.

Efficient sampling of object surface space and viewpoint space is achievable due to the high degree of intrinsic redundancy. A fairly coarse rough model mesh will suffice for view planning at the fine modeling stage, provided the rough model correctly captures the object's gross topology and geometry. Secondly, complex shape 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 measurability estimation on a coarse object surface representation with respect to a modest number of well-chosen candidate viewpoints. In pseudo code form, the 3M algorithm proceeds as follows at the fine modeling stage:

```
Decimate the rough mesh model to a suitably coarse representation
Segment the rough model by surface feature type (optional)
While (Unscanned rough model regions)
  Generate optimal candidate viewpoint set
  Compute measurability; form measurability matrix
  Solve the set covering problem; Select NBV(s)
```

Table 3.1: 3M Algorithm Pseudo Code

Rough Model Decimation The rough model is decimated to a level just adequate for view planning purposes. This level is experimentally determined. Chapter 5 examines issues relating to sampling object surface space.

Rough Model Segmentation Optionally, the rough mesh model is segmented into its major shape components such as cavities, holes, protrusions and planar patches. We will often refer to segments as surface patches. Segmentation has the advantage of reducing the net cost of computing measurability data for the object[1]. Its disadvantages include the following. There

[1] Refer to Section 3.6 for a discussion of computational complexity.

is the cost of an additional preprocessing step. On average, composite view plans will be slightly longer due to patch edge effects. Rework may be necessary as a result of occlusion errors arising from view planning being localized to major surface patches. The decision to use or not use segmentation will depend on the application and the processing power available for the view planning function.

Viewpoint Generation For each surface patch (or the entire rough model if not segmented), a set of candidate viewpoints is generated that is most likely to measure the surface in accordance with the specified criteria. Viewpoint generation is achieved by quantizing the optimal viewpoint zone for the specific sensor and surface feature geometry. Chapter 6 examines schemes for efficiently sampling viewpoint space. Here, we will briefly describe one viewpoint generation scheme, the "optimal scanning zone algorithm".

An ideal candidate generalized viewpoint $(\mathbf{v}, \lambda_s)$ is created for each surface point - that is, for each vertex of the rough mesh model. It is convenient to consider pose orientation as the concatenation of two rotations - boresight axis and rotation about the boresight axis (twist). Thus, for view planning purposes, we use pose = (position, axis, twist). Internally, homogeneous transforms are used where appropriate.

To generate a candidate viewpoint, each rough model vertex is dilated along the local surface normal by a stand-off distance adjusted to $f_d R_o$ where R_o is the optimal stand-off distance. The stand-off adjustment factor f_d is set to $f_d = 1 + \delta$, where δ is an experimentally-derived constant $\ll 1$, to allow for errors in the estimation process and positioning system. The dilated surface vertex becomes the position of the corresponding candidate viewpoint. The viewpoint's axis is set to the reverse of the dilation vector. Setting the sensor twist orientation is less straight forward as it is influenced by shadow effects arising from object shape over a broader surface region. As an initial approach pending further discussion, we set the twist angle to minimize local shadow effects. Finally, configurable sensor parameters are optimized for each candidate viewpoint, based on the model specification, system capabilities and observable object characteristics.

We postulate that the set of viewpoints thus generated provides a near-optimal sampling of viewpoint space for the purposes of solving the view planning problem. As viewpoint parameters are optimized on a per-viewpoint basis, this discretization scheme reduces the dimensionality of viewpoint space from $6D^+$ to 2D. View planning optimization is conducted over the domain of generalized viewpoints.

Measurability Estimation Using the model specification and performance models for the sensor and positioning system, a measurability matrix is computed for the entire rough model or, if segmentation has been used, a measurability matrix is computed for each surface 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. Each measurability computation is required to pass the frustum occupancy, visibility and specification-compliance tests previously described. For segmented models, measurability determination at this stage is based on local visibility analysis only - that is, considering only the geometry of the target surface patch.

NBV Selection Upon computation of the measurability matrix or matrices, a number of techniques are available to solve the set covering problem. The output is a NBV set or view plan. If the rough model has been segmented, the component view plans are merged. Set covering algorithms are

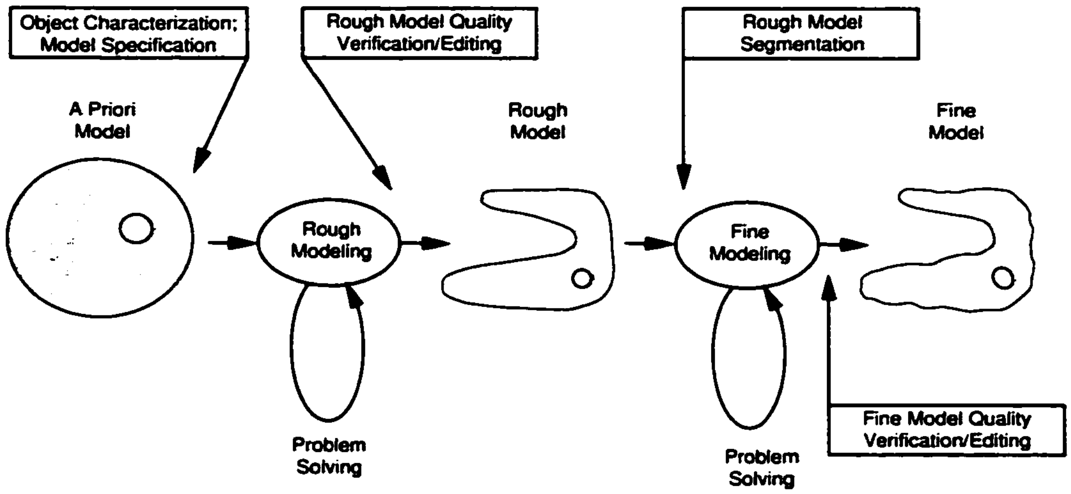


Figure 3.8: Human Interaction with Reconstruction Process

discussed in the next chapter.

Operational Concept At the current state of research, the 3M view planning concept envisions a semi-automated object reconstruction process. Apart from physical tasks such as placing and re-positioning the object within the imaging workspace and other set-up tasks, the operator will be expected to perform various processing tasks as illustrated in Figure 3.8. The operator is responsible for input task specification and output quality verification and possible editing. Additionally, the operator will be required to verify the output of the rough modeling stage and make corrections, if necessary.

3.5.2 Clarification of Thesis Scope

This thesis addresses view planning for the rough model to fine model stage of model-based object reconstruction (Figure 3.3). It excludes the rough modeling stage, automated segmentation and automated fine model problem resolution. We believe automated or semi-automated rough modeling is feasible by means of a rapid, pre-programmed view plan[2] coupled with robust model building techniques [RW97]. Automation of the segmentation task is considered a feasible but non-trivial problem. Where segmentation is used in this thesis work, it is done manually. Automated or semi-automated fine model problem solving is considered to be less of a challenge and to involve a mix of development and research. The research presented in this thesis addresses all of the view planning requirements derived at Section 1.2 with the exception of the material constraints requirement, some types of measurement artifact and one aspect of the overlap constraint. We assume objects have benign reflectance characteristics i.e. reasonably uniform, moderate reflectance characterized by Lambertian scattering. The issue of view planning for more difficult reflectance characteristics, in particular the specular reflectance case, is briefly addressed in the final chapter under future work.

3.5.3 Comparison with Previous Work

The modified measurability matrix view planning algorithm presented in this thesis involves modifications and extensions to the original measurability matrix work by Tarbox. Besides being applied to object reconstruction rather than inspection, the key points distinguishing our work from Tarbox

[2]The approach to automated rough modeling will depend on the imaging environment - that is, the relative object size and the capabilities of the sensor and positioning system. A view sphere approach (Chapter 6) would be suitable for objects whose size is comparable to or smaller than the sensor frustum. For large objects, a space carving technique would be suitable (Chapter 2).

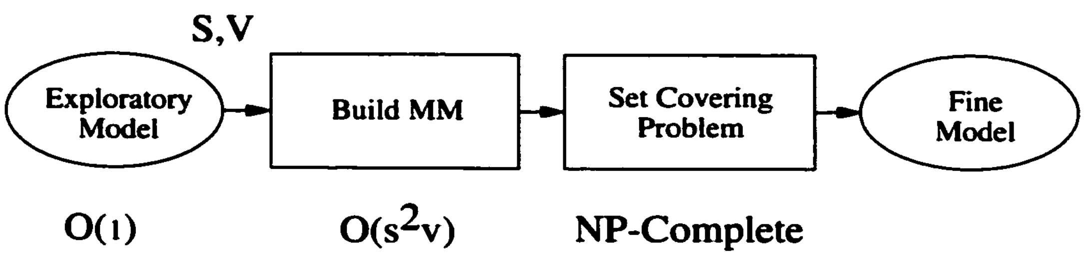


Figure 3.9: View Planning Problem - Model-Based Perspective

are the following. Our method is performance-oriented - that is, it is based on a model specification. We use more stringent and objective measurability criteria based on sensor and positioning system performance models. Computational complexity is reduced by optimized, sparse sampling in S and V as well as optional segmentation of the problem into smaller components.

In the concluding chapter, we will evaluate this approach with respect to the view planning requirements derived in the introduction.

3.6 Computational Complexity

Figure 3.9 presents an overview of the complexity of the view planning problem from a model-based perspective. The process of physically scanning and building the exploratory model will take time, of course. However, as a preplanned activity, it has constant complexity. The set covering problem is NP-complete, as previously noted, but there are a variety of solution techniques, both exact and approximate, as we will discuss in the next chapter. The computational crux of the model-based approach to view planning is constructing the measurability matrix or matrices.

From the structure of the measurability matrix, we can immediately determine its computational complexity. $\mathbf{M}$ has sv elements, where $s = |S|$ is the number of rows (surface points) and $v = |V|$ is the number of columns (viewpoints). The most computationally intensive part of computing a measurability matrix is visibility analysis. In general, a ray tracing operation must be performed twice per measurability matrix element - once each for the optical source and receiver[3]. Each ray tracing operation involves testing transmit or receive rays with each triangle t_k of the rough model mesh. As the number of triangles in the mesh $t \approx s$, then the computation complexity of deriving a measurability matrix is approximately $\mathcal{O}(s^2v)$.

If S is segmented into c segments of roughly equal size, then s and v are proportionately reduced. In this case, the overall computational complexity of computing c measurability matrices is $\mathcal{O}(c\frac{s^2}{c^2}\frac{v}{c}) = \mathcal{O}(\frac{s^2v}{c^2})$.

It is useful to quantify these variables in the context of a brute force discretization of S and V. The size of a typical reconstruction is $s \approx 10^5$. Discretizing a one cubic meter imaging workspace at one centimeter intervals gives 10^6 sensor positions. After discarding infeasible orientations, quantizing rotation in five degree increments produces approximately 10^5 sensor orientations. In combination, this generates approximately 10^{11} sensor poses. To compute the total number of viewpoints, we would have to add a further quantization of say ten levels for each configurable sensor parameter. So, if this viewpoint generation scheme were followed, the number of viewpoints

[3]Tarbox reduced visibility tests from one per source and receiver position to one per view sphere location by a clever discretization scheme. A view sphere was approximated by a recursively sub-divided icosahedron and viewpoints were defined as the set of all nodes on the view sphere separated by a distance equal to the optical baseline. The gain in computational efficiency was achieved at the expense of a fixed and coarse discretization of the rotation component of sensor pose for a fixed view sphere radius.

would be approximately $v \approx 10^{11+}$. The resulting number of visibility calculations for a brute force quantization of object surface space and viewpoint space would be of the order of 10^{21+}. Consequently, it is readily apparent that success in view planning with measurability matrices is dependent on suitable schemes for sparse sampling in object surface space and viewpoint space. At the same time, the power of the technique is also immediately apparent for we have, in one data structure, all the information necessary to construct an accurate, robust and efficient view plan.

This examination of basic computational complexity issues raises some important questions:

How rough can the rough model be for effective view planning? That is, roughness in terms of sampling density and sampling errors.

How tolerant is the view plan to pose uncertainty?

What is the optimal scheme for sampling viewpoint space?

We will address these questions in the coming chapters. First, however, we will present a theoretical framework for the view planning problem.

Chapter 4

View Planning as a Set Covering Problem

This chapter provides a theoretical framework for the view planning problem (VPP) in the form of a set covering problem (SCP) expressed as an integer programming problem (IP). We also present experimental results for a greedy search set covering algorithm modified to include a registration constraint.

4.1 Theoretical Background

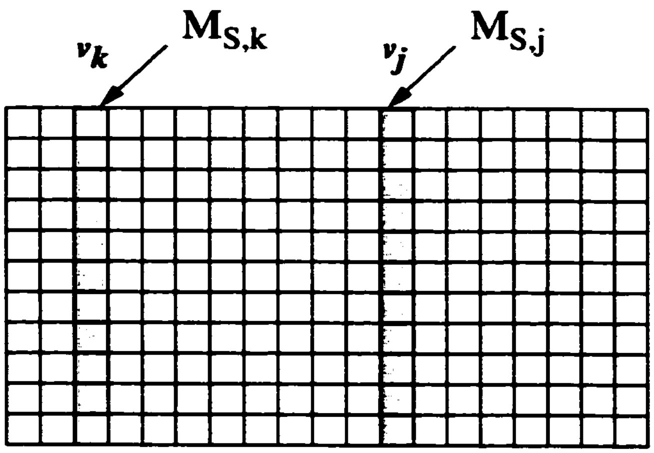
To date, theoretical treatment of the view planning problem has been limited. Tarbox and Gottschlich [TG95] introduced the measurability matrix concept in a model-based approach to inspection. They showed that the VPP was isomorphic to the set covering problem which is known to be NP-complete. Yuan [Yua95] used mass vector chains in a global approach to view planning. Whaite and Ferrie [WF97] presented an autonomous exploration theory based on minimization of model uncertainty. They observed that the best sensor locations are those where the ability to predict is worst - that is, where the variance in the fit of the data to the current model is greatest.

Soucy et al [SCF98] examined view planning from the perspective of an infinitely small surface explorer and local topological operations on a voxelized surface. They also took the capabilities and limitations of the positioning system into account. Arbel and Ferrie [AF99] presented an entropy-based gaze planning strategy in the somewhat related field of object recognition. Scott et al [SRR00] extended Tarbox's work to object reconstruction with a set theory based formulation of the VPP in terms of a measurability mapping between viewpoint space and object surface space. Most other work in the field relies on a variety of heuristic techniques without a theoretical framework.

The referenced work has contributed to a better understanding of the view planning problem, yet the field lacks a comprehensive theoretical foundation. A truly comprehensive theory would encompass all problem constraints and error mechanisms and provide a theoretical basis for sampling surface and viewpoint space. In the following, we define a theoretical framework dealing with measurability and registration constraints. We begin by developing a registration constraint suitable for an integer programming formulation. We then apply the theoretical formulation to experimentation with a greedy search set covering algorithm modified to include a registration constraint.

4.2 Image-Based Registration Constraint

A common imaging problem arises when the precision of the positioning system is inferior to that of the sensor. It is then necessary to employ image-based registration to bring images into a common reference frame with a precision comparable to that of individual surface measurements. The iterative closest point (ICP) algorithm [BM92] with its many recent enhancements is widely used for this purpose. The algorithm works by minimizing errors in



$$\sigma_{kj} = \frac{M_{S,k} \bullet M_{S,j}}{m_s}$$

$$m_s = \max \| M_{S,k} \|$$

Figure 4.1: Viewpoint Correlation

the geometric fit between overlapping range image segments. In this section we show how a range image-based registration constraint can be expressed within an integer programming formulation of the VPP.

4.2.1 Viewpoint Correlation

Image-based registration requires sufficient overlap between images[1]. A degree of image overlap is also necessary for image integration - the process of building a single, non-redundant mesh model from multiple range images. As a first approximation, this initial algorithm specifies a point overlap con-

[1]Few view planning techniques in the literature incorporate a registration constraint. Pito [Pit96b] includes an explicit overlap requirement and mentions the need for shape complexity in the overlap area but does not implement it. Whaite and Ferrie [WF97] achieve image overlap by a conservative search strategy.

straint which is necessary but not sufficient for image-based registration[2].

There is a direct correspondence between viewpoints and range images. Image overlap can therefore be determined from the degree of viewpoint correlation. For view planning purposes, we define the cross-correlation σ_{kj} of viewpoints v_k and v_j as the inner product of the respective column vectors $\mathbf{M}_{S,k}$ and $\mathbf{M}_{S,j}$ of the measurability matrix (Figure 4.1), normalized[3] by the maximum surface coverage of any viewpoint in the candidate viewpoint set - that is, $m_S = max||\mathbf{M}_{S,k}||\ \forall k \in V$, so

$$\sigma_{kj} = \frac{\mathbf{M}_{S,k} \cdot \mathbf{M}_{S,j}}{m_S}. \tag{4.1}$$

To register image (viewpoint) v_k with image (viewpoint) v_j with sufficient overlap, we require that their cross-correlation σ_{kj} exceed a registration threshold t_r, typically around 20%, i.e.

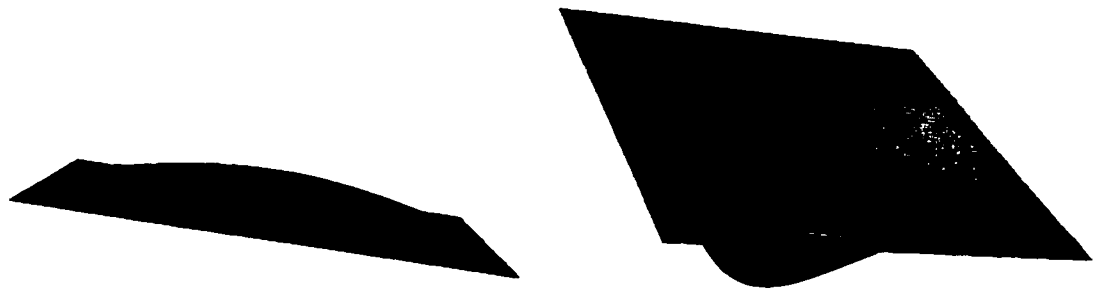
$$\sigma_{kj} \geq t_r. \tag{4.2}$$

Specifying a value for the registration threshold t_r is equivalent to specifying that two range images must overlap by that amount to satisfactorily register with each other.

---

[2]In general, we need to add a geometric complexity requirement to the overlap region to fully constrain registration in all translations and rotations. It will be apparent from the simpler point overlap constraint how, in principle, we can formulate a more stringent constraint for overlap with geometric complexity.

[3]There are several potential choices for a normalizing value. As defined, m_S has the advantage of guaranteeing σ_{kj} values in the range [0,1], is reciprocal and is independent of segmentation patch size, object size, rough model sampling density and sensor characteristics.



Slightly Convex Patch Deeply Concave Patch

Figure 4.2: Viewpoint Correlation Test Patches

4.2.2 Footprint Ratio

The problem of covering a surface patch can be characterized by the size and shape of the patch relative to sensor coverage. Convenient shape measures include depth-to-width ratio D/W for convex and concave patches. Relative size can be characterized by the following measures: the *patch footprint* F_p equal to the patch area and the *sensor footprint* F_s equal to the frustum cross-sectional area at the sensor's optimal scanning range R_o. A surface patch's *footprint ratio* $r_f = F_p/F_s$ provides a rough indication of the match between patch size, sensor capability and number of views required. The absolute lower bound on the number of views required to cover a patch is $\lceil r_f \rceil$. Practically speaking, the lower bound is attainable only under the most optimistic scenario of a beneficial match between the shape of the patch and the sensor footprint.

4.2.3 The VPP Solution Landscape

From our human visual experience, we intuitively understand the phenomenon of viewpoint redundancy. In the context of view planning, it is useful to quantify and visualize viewpoint correlation with some simple examples. With this objective in mind, we generated two synthetic surface patches, the first larger than the sensor footprint and slightly convex, and the second smaller than the sensor footprint and deeply concave (Figure 4.2). Surface noise was not added for these experiments. Viewpoint space was sampled by generating a set of optimal candidate viewpoints, one per mesh vertex, following the "optimal scanning zone algorithm" described in the previous chapter.

Figures 4.3 and 4.4 show viewpoint correlation relative to the center vertex of each test patch as a function of translation on the optimal scanning zone in viewpoint space. In these figures, u and v are translation parameters for viewpoints on the optimal scanning surface. The axis of correlation symmetry corresponds to the sensor optical baseline. Correlation is not symmetric along the baseline as measurability is not reciprocal with respect to laser and detector locations. The data indicates that neighbouring viewpoints are highly correlated, as expected, and that correlation falls off with displacement and is modified by surface shape. The footprint ratios are 2.55 and 0.55 respectively, while the measurability matrix densities are 0.238 and 0.637.

Thus, in contrast to many set covering applications (SCP) reported in the literature [GW97], variables in the view planning set covering problem exhibit a high degree of correlation within viewpoint space neighbourhoods. Additionally, we note that the density of the constraint matrix (measurability matrix in our case) is medium to high whereas it is sparse in many traditional SCP applications. Finally, as in standard SCP applications, VPP solutions

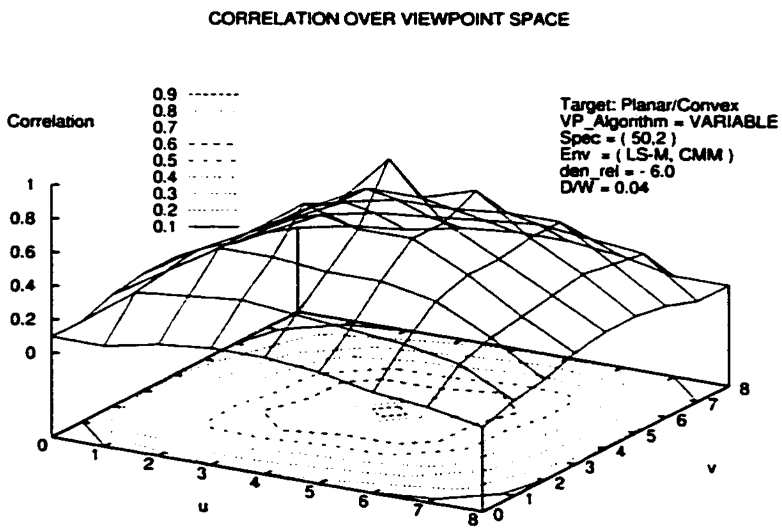


Figure 4.3: Viewpoint Correlation - Planar/Convex

are generally not unique.

A VPP set covering algorithm must contend with competing forces acting on viewpoints in a candidate NBV set - a repelling force minimizing viewpoint correlation to gain coverage and an attracting force seeking to achieve a minimum correlation level to maintain compliance with the registration constraint. A physics-based VPP set covering algorithm may be possible based on this paradigm.

4.2.4 Registration Constraint

Viewpoint correlation, redundancy and overlap are essentially the same thing. Pose uncertainty resulting from positioning system inaccuracy results in a requirement for image-based registration which in turn requires a specified

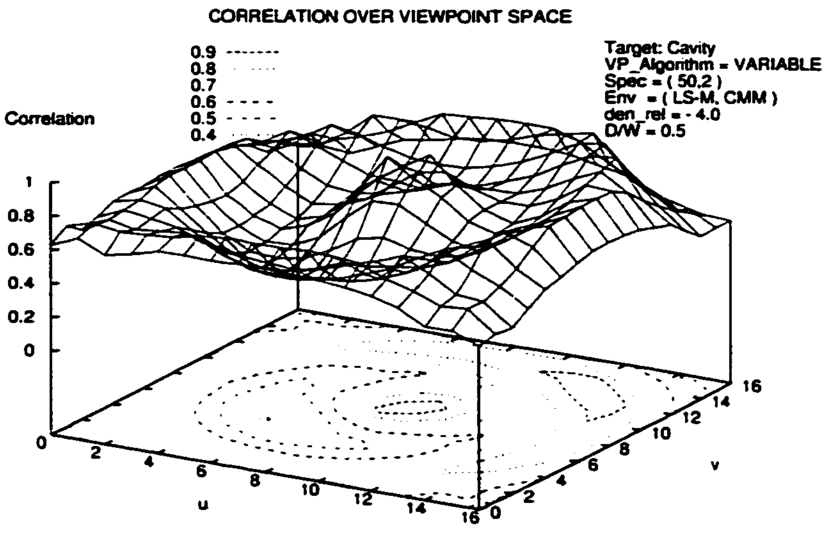


Figure 4.4: Viewpoint Correlation - Cavity

degree of image overlap i.e. viewpoint correlation. Viewpoint correlation has another consequence. It impacts the discretization scheme for viewpoint space. We wish to have a set of candidate viewpoints sufficiently nearby to satisfy registration and integration requirements yet far enough apart for efficient sampling of viewpoint space. Understanding viewpoint correlation in our problem set will assist in optimizing these competing objectives.

Further, we can observe that an overlapping range image set has an associated registration graph G_r (Figure 4.5) which encodes viewpoint connectivity in terms of inter-image registration potential. Two nodes (viewpoints) in a registration graph are connected if the range images associated with those viewpoints have sufficient overlap to be registered with each other using an image-based registration algorithm such as the Iterative Closest Point (ICP)

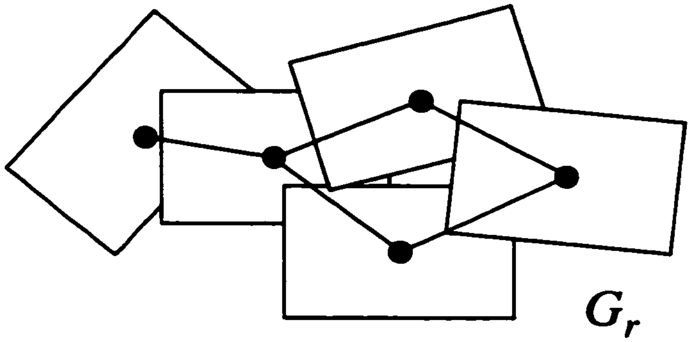


Figure 4.5: Image-Based Registration Graph G_r

algorithm. Consequently, we can express a basic image-based registration requirement for view planning by stating that

The viewpoint registration graph must be at least simply connected.

With segmentation, the viewpoint registration graph for each patch is a sub-graph of the global registration graph for the entire object. We can ensure global connectivity at the object level by an initial segmentation with suitable overlap across segmentation boundaries.

4.2.5 Constructing the Registration Graph

Given a measurability matrix $\mathbf{M}$, the issue then is first how to create the registration graph corresponding to any viewpoint set $N \subset V$ and, secondly, how to utilize the connectivity information embedded in the graph to express a set of registration constraints on the view planning problem.

A process for achieving this is illustrated in Figure 4.6. We begin by computing the symmetric viewpoint *correlation matrix* $\Sigma = [\sigma_{kj}]$. Σ specifies viewpoint adjacency in registration terms. We then define a viewpoint

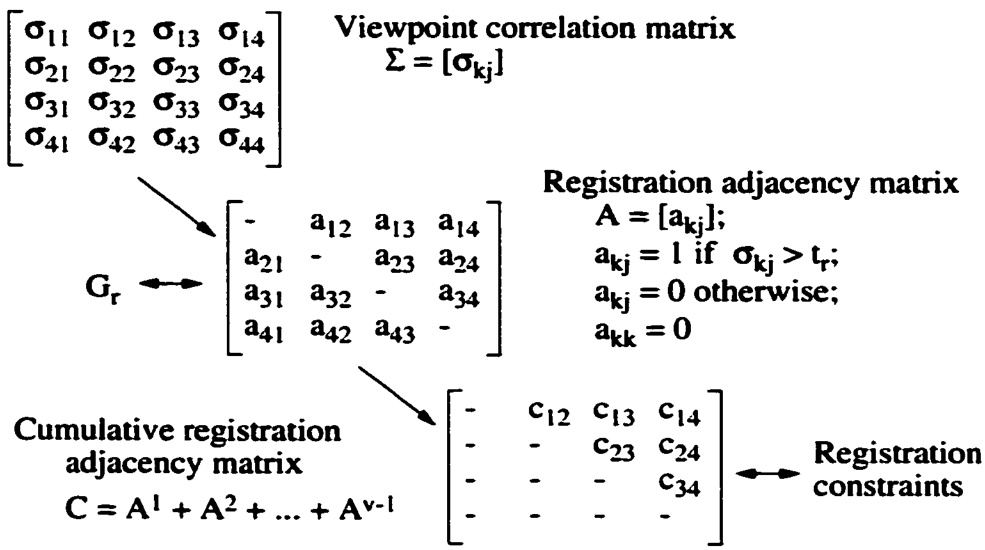


Figure 4.6: Registration Constraint Creation

registration adjacency matrix $A = [a_{kj}]$ such that $a_{kj} = 1$ if $\sigma_{kj} \geq t_r$, $k \neq j$, and $a_{kk} = 0$. The registration adjacency matrix A has an associated *viewpoint registration graph* G_r.

Connectivity information embedded in the registration-adjacency matrix can be used to determine registration graph connectivity. The $(k, j)^{th}$ entry of the matrix product A^r defines the number of different paths (including backtracking) between nodes k and j of length r in the graph. The maximum length of a non-repeating path in the registration graph is one less than the size v of the viewpoint set V. Consequently, the *cumulative registration adjacency matrix* $C = [c_{kj}]$ defined as follows[4] captures the number of paths

[4]Thanks to Professor P.K. Allen for observing that C is equivalent to the transitive closure of G_r.

of all lengths from minimum to maximum connectivity between viewpoints:

$$C = A^1 + A^2 + \cdots + A^{v-1}. \tag{4.3}$$

C can be computed directly from Σ as a preprocessing step prior to solving the set covering problem. Further, C can be computed numerically or symbolically. The advantage of symbolic computation (illustrated by a simple example at Appendix B) is that it can be computed once for $\mathbf{M}$ and then applied to any view plan $N \subset V$.

For the registration graph to be connected, elements in rows and columns of C corresponding to viewpoints in the next-best-view set N must be at least one. Therefore, a necessary condition for successful image-based view registration of two viewpoints is

$$c_{kj} \geq 1\,;\; k, j \in N\,; k \neq j. \tag{4.4}$$

As C is symmetric with all elements ≥ 0, and diagonal elements are irrelevant, the equation can be simplified as follows, where $n = |N|$ and, generally, $n \ll v$:

$$c_{kj} \geq 1\,;\; k = 1, \cdots, (n-1)\,;\; j = k+1, \cdots, n\,;\; k, j \in N. \tag{4.5}$$

Furthermore, C defines registration connectivity of all viewpoints in V, whereas we wish to apply the global registration constraint only on $N \subset V$. Thus, we need to define $\mathbf{X}$ as a vector of binary viewpoint variables x_j such that $x_j = 1$ if column j is in the solution and $x_j = 0$ if it is not. Then, if $x_l = 0$, ignore c_{kj} for all $k, j = l$. We can achieve the desired effect over the full problem domain by changing the registration constraint to

$$c_{kj} \geq x_k x_j\,;\; k = 1, \cdots, (v-1)\,;\; j = k+1, \cdots, v\,;\; k, j \in V \tag{4.6}$$

We can now proceed to integrate the registration constraint in an overall integer programming formulation of the view planning problem.

4.3 Integer Programming Formulation of the View Planning Problem

4.3.1 IP Formulation

In the previous chapter, we showed why the view planning problem is isomorphic to the set covering problem (SCP). Recognizing the VPP as an instance of the SCP admits its expression as a classical 1/0 integer programming problem (IP), a sub-class of linear programming problems (LP). The view planning problem can then be expressed as the problem of covering the rows of a binary s-by-v measurability matrix $\mathbf{M} = [m_{ij}]$ by a minimal subset of its columns. Recall that rows of a measurability matrix correspond to surface points while columns correspond to viewpoints.

Informally, the goal of view planning is to find a suitably small set of views that cover all the surface points while satisfying the measurability constraints. More formally, the problem can be written as follows, where notation follows integer programming conventions:

$$\text{Minimize } Z = \sum_{j=1}^{v} c_j x_j \tag{4.7}$$

subject to

$$\sum_{j=1}^{v} m_{ij} x_j \geq 1\,;\; i = 1, \ldots, s\,;\; i \in S, \tag{4.8}$$

$$c_{kj} \geq x_k x_j\,;\; k = 1, \cdots, (v-1)\,;\; j = k+1, \cdots, v\,;\; k, j \in V, \mathit{and} \tag{4.9}$$

$$x_j \in \{0, 1\}\,;\ j = 1, \ldots, v\,;\ j \in V. \tag{4.10}$$

Equation 4.10 applies an integer constraint on the viewpoint variable. $\mathbf{X} = [x_j]$ spans viewpoint space V as sampled by the viewpoint generation stage. The optimal $\mathbf{X}$ is the minimal set of viewpoints covering object surface space S - that is, the next-best-view set N. In the objective function, equation 4.7, coefficients c_j allow a movement cost to be associated with each viewpoint. This would be appropriate, for example, in the case of a positioning system with significant, non-uniform movement costs[5]. Otherwise, the costs can be set to 1, a formulation known as the unicost set covering problem [BC96], a somewhat easier problem to solve. The constraints defined by equation 4.8 ensure that each row of the measurability matrix (i.e. each surface point) is covered by at least one viewpoint. Equation 4.9 imposes the image-based registration constraint developed in the previous section. The range of indices in equation 4.9 indicates we are interested only in the upper right triangle of the cumulative registration adjacency matrix.

Viewpoint feasibility is determined by the range of motion and degrees of freedom of the positioning system. There is no need to include a viewpoint feasibility constraint in the above IP formulation. It is more efficient to avoid generating infeasible viewpoints at the viewpoint generation stage.

4.3.2 Set Covering Algorithms

Expressing the view planning task as an IP provides a compact mathematical formulation of the problem, opening the application to the rich research base

[5]An example would be the case of a rotary joint limited to a discrete set of accurate orientations where change requires a unlock-rotate-relock sequence. In some cases, system recalibration may be required upon reconfiguration of a positioning system component.

in discrete combinatorial optimization. Unfortunately, in general an integer programming problem is considerably more difficult to solve than the equivalent LP problem. IP solution time can be highly unpredictable, depending on problem formulation, data characteristics and problem size (number of variables and number of constraints). Optimal solution methods such as branch-and-bound and cutting-plane techniques [Win95] typically use an intelligent tree search of feasible solutions and are found in a variety of commercial LP/IP solvers. While guaranteeing optimal results, such exact methods can be computationally prohibitive even for modestly sized IPs. For most medium-to-large IPs, this leaves a choice of approximate and heuristic algorithms [Ree93], including greedy search (GS) [FW82], simulated annealing (SN) [Sen93], genetic algorithm (GA) [BC96], Lagrangian relaxation [Bea90] and neural network [GW97] methods. Most published performance results [GW97], [Bea90] deal with random, low density data sets. The VPP falls into the category of a medium-to-large IP with non-random data and moderate density. A representative size for a measurability matrix generated by our algorithm for a segmented problem will be approximately 400 variables by an equal number of constraints, or about $2x10^5$ elements.

4.3.3 Discussion of the IP Formulation

The foregoing theoretical framework for the view planning problem cast in the form of a set covering problem and expressed as an integer programming problem is beneficial from several aspects. It provides more comprehensive theoretical treatment of the performance objectives, constraints and error mechanisms associated with view planning than heretofore. For the first time, it provides an explicit solution mechanism using either exact or approximate integer programming techniques. Integer programming is a well-understood field, as is the set covering problem. Framing the view planning problem in terms of well-known mathematical problems and techniques opens up a large body of existing knowledge in these fields. Finally, the straight forward mathematical formulation assists with a more intuitive understanding of the nature of the problem.

A truly comprehensive theory would encompass sampling schemes for surface and viewpoint space, and all problem constraints and error mechanisms. In relation to this goal, our theoretical formulation addresses the measurability constraint and the overlap component of the registration constraint. Surface sampling error is not directly addressed but is dealt with separately in the next chapter. As will be discussed in Chapters 6 and 7, pose error is accommodated at the viewpoint generation stage. Non-uniform movement costs associated with the positioning system are incorporated directly in the objective function. Constraints in the degrees of freedom and range of motion of the positioning system are dealt with at the viewpoint generation stage. Surface and viewpoint space sampling is dealt with empirically in Chapters 5 and 6, respectively. Chapter 6 provides a basic theoretical framework for viewpoint space sampling. A theoretical basis for surface space sampling remains an open problem.

Some observations on the registration constraint are appropriate. Firstly, we note the constraints are non-linear, making the IP non-linear. However, we can solve for the set covering constraints alone and then enforce registration compliance as part of the fitness function. If required, a minimal number of views can be added to satisfy the registration constraint and render the solution feasible overall.

A second concern is the computational cost of formulating the registration constraint as calculation of C involves symbolic matrix operations on large matrices. The computational complexity of C is approximately $\mathcal{O}(v^4)$

operations, where v is the number of viewpoints in V. However, the matrix need be computed only once. Additionally, we have the option to delay computation of C to only candidate solutions (i.e. candidate NBV sets), in which case the dimensionality of the required computations is greatly reduced and can be numerical rather than symbolic. In the latter case, the computational complexity using the cumulative registration adjacency matrix technique[6] is approximately $\mathcal{O}(cn^4)$ operations, where c is the number of candidate solutions, n is the size of NBV set N and $n \ll v$.

In summary, the formulation of the view planning problem as an integer programming problem provides an explicit solution mechanism in the context of classical mathematical problems and solution techniques. It also aids in a more intuitive understanding of the nature of the problem.

4.4 Greedy Search Set Covering

In this section, we show how the integer programming formulation of the view planning problem can be implemented by a fast and robust approximate set covering algorithm.

4.4.1 Experimental Process

The experimental process followed in this work (Figure 4.7) employs the steps of the modified measurability matrix algorithm (Figure 3.3) in a closed loop process with synthetic data to explore the impact of system parameters,

[6] However, there are more efficient connectivity algorithms for a fixed graph structure. For example, Moore's breadth first search algorithm [Jun99] has complexity $\mathcal{O}(|V||E|)$ where $|E|$ is the number of edges in the graph. Other standard transitive closure algorithms include Floyd-Warshall and Dijkstra [Jun99].

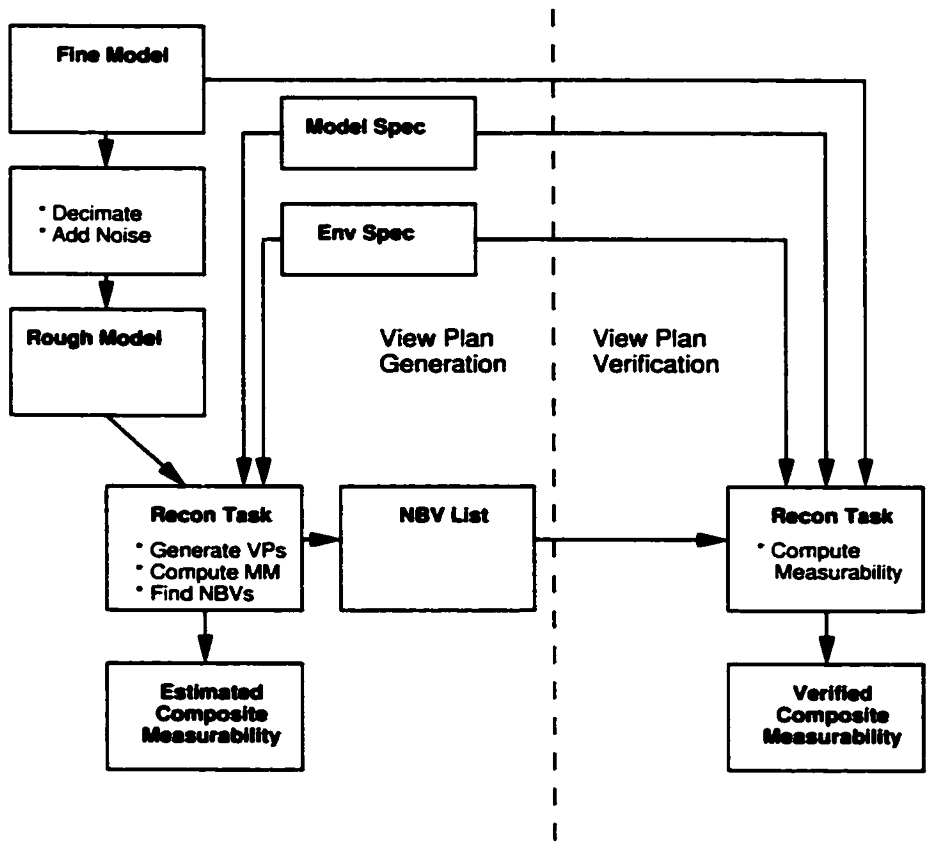
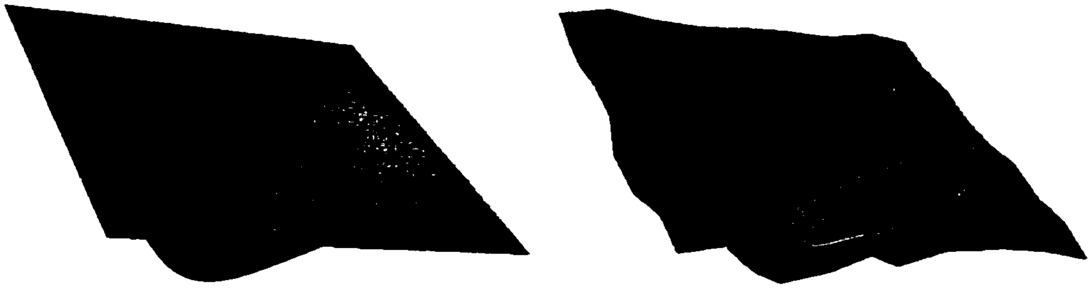


Figure 4.7: Experimental Process - Measurability Verification

system errors and object shape on view plan quality. Synthetic shapes are used primarily because they are repeatable, quantifiable and scalable. A high density mesh such as the example at the left in Figure 4.8 serves as a surrogate fine model. The fine model is decimated to create a rough model at a desired lower sampling density, after which surface noise may be added. An example coarsely-sampled, noisy rough model is shown at the right in Figure 4.8. The rough model drives the view planning process, including



Cavity Fine Model Noisy Cavity Rough Model

Figure 4.8: Cavity Models

viewpoint generation, measurability estimation and NBV set determination. Other inputs include imaging environment and model specifications. The former defines performance models for the sensor and positioning system. The latter specifies model quality objectives - presently in the form of the tuple (measurement precision, sampling density). The resulting view plans are executed against the parent fine model, closing the loop to provide verified measurability. Measurability terms are defined in Section 3.3.2.

4.4.2 Greedy Search

In this section we show experimental results using the greedy search (GS) set covering algorithm. The standard GS algorithm begins by computing a figure of merit (FOM) for each non-selected viewpoint - that is, the ratio of currently uncovered surface points covered by a candidate viewpoint to the total currently uncovered. Then, greedy search simply selects the viewpoint with the highest FOM.

The first experiment was conducted with a nearly planar patch ($D/W =$

0.04) whose area was moderately greater than the sensor footprint (r_f = 2.55). The absolute lower bound on the number of views required to cover this patch is $\lceil r_f \rceil = 3$. Visual inspection shows the real lower bound to be 4, given the actual shape of the patch and sensor footprint. The experiment was conducted at a low rough model relative sampling density $\rho_{rel} = -6.0$ and a model specification of (50 μm, 2 s/mm^2) for a range camera whose best-case performance was approximately (10 μm, 10 s/mm^2). Relative sampling density ρ_{rel} is defined as the logarithm to the base two of the ratio between the sampling density of the rough and fine models. This case illustrates a set covering problem driven more by relative size than shape.

Figure 4.9 compares the progression of estimated measurability by the GS algorithm with the optimal solution obtained by exhaustive search (ES). The optimal curve represents the best achievable for any given size i of next-best-view set N_i. We note that greedy search results compare favourably to the optimal achievable with the current level of viewpoint space discretization (seven views required versus six), while the computational complexities are $\mathcal{O}(v^2)$ and exponential, respectively[7].

Figure 4.10 presents closed loop experimental results by verifying the computed NBV set against the fine model. Estimated measurability is very close to the verified performance, despite the low relative sampling density. The verified measurability in this case is $m_v = 0.9965$.

The NBV sequence generated by the GS algorithm is shown in Figure 4.13 on a parameterized plot of the optimal scanning surface in viewpoint space. In Figures 4.13 and 4.14, u and v are translation parameters for viewpoints

[7]As will be discussed later, a more efficient discretization of viewpoint space will provide a true optimum view plan length of four. Our actual view plan efficiency, then, is $e_v = 4/7 = 0.57$.

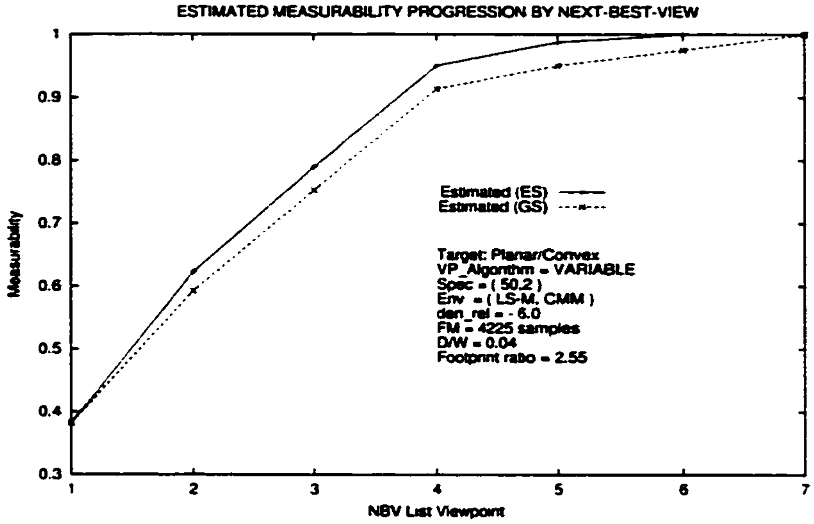


Figure 4.9: Estimated Measurability Progression

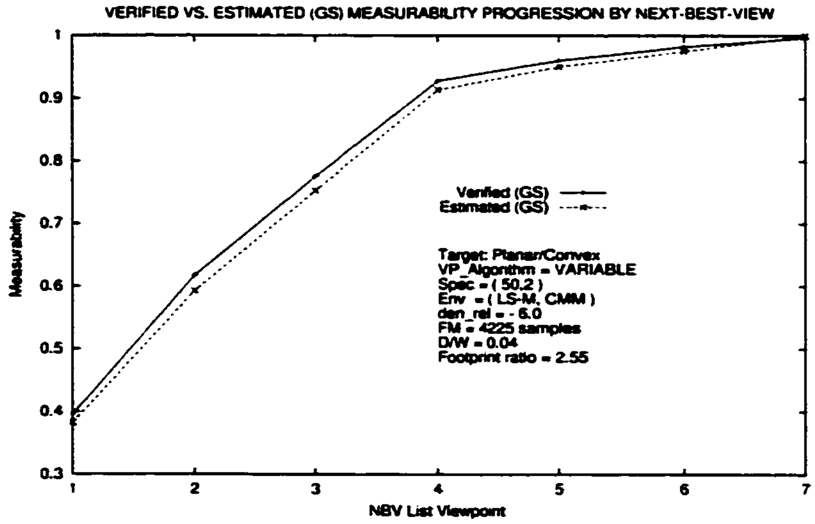


Figure 4.10: Verified Versus Estimated Measurability Progression

on the optimal scanning surface. These parameters have a direct correspondence to translation on the surface patch in object surface space. The GS set covering can be compared to the optimal set covering shown in Figure 4.14. The first six of the sequence of seven next-best-views for the GS set covering are shown in Figures 4.11 and 4.12. Individual shaded measurability projections are shown in the left column in sequence; composite measurability projections on the right. The shading is white for measurable, black for non-measurable. Other shading is an artifact of the rendering program.

The matrix of raw correlation values for the GS next-best-view set at Figures 4.9, 4.10 and 4.13 is as follows.

$$\begin{bmatrix} - & 0.326 & 0.097 & 0.097 & 0.326 & 0.226 & 0.387 \\ 0.326 & - & 0 & 0 & 0 & 0.194 & 0.355 \\ 0.097 & 0 & - & 0 & 0 & 0.452 & 0 \\ 0.097 & 0 & 0 & - & 0 & 0 & 0.326 \\ 0.326 & 0 & 0 & 0 & - & 0 & 0 \\ 0.226 & 0.194 & 0.452 & 0 & 0 & - & 0.065 \\ 0.387 & 0.355 & 0 & 0.326 & 0 & 0.065 & - \end{bmatrix}$$

From it we can deduce the connectivity of the registration graph G_r if various levels of registration threshold t_r were applied post-facto. For example, for $t_r = 0.2$, the registration adjacency matrix A is as follows. Inspection of A shows that the associated registration graph G_r is connected - in this case by statistical accident rather than by design.

$$\begin{bmatrix} - & 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & - & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & - & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & - & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & - & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & - & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & - \end{bmatrix}$$

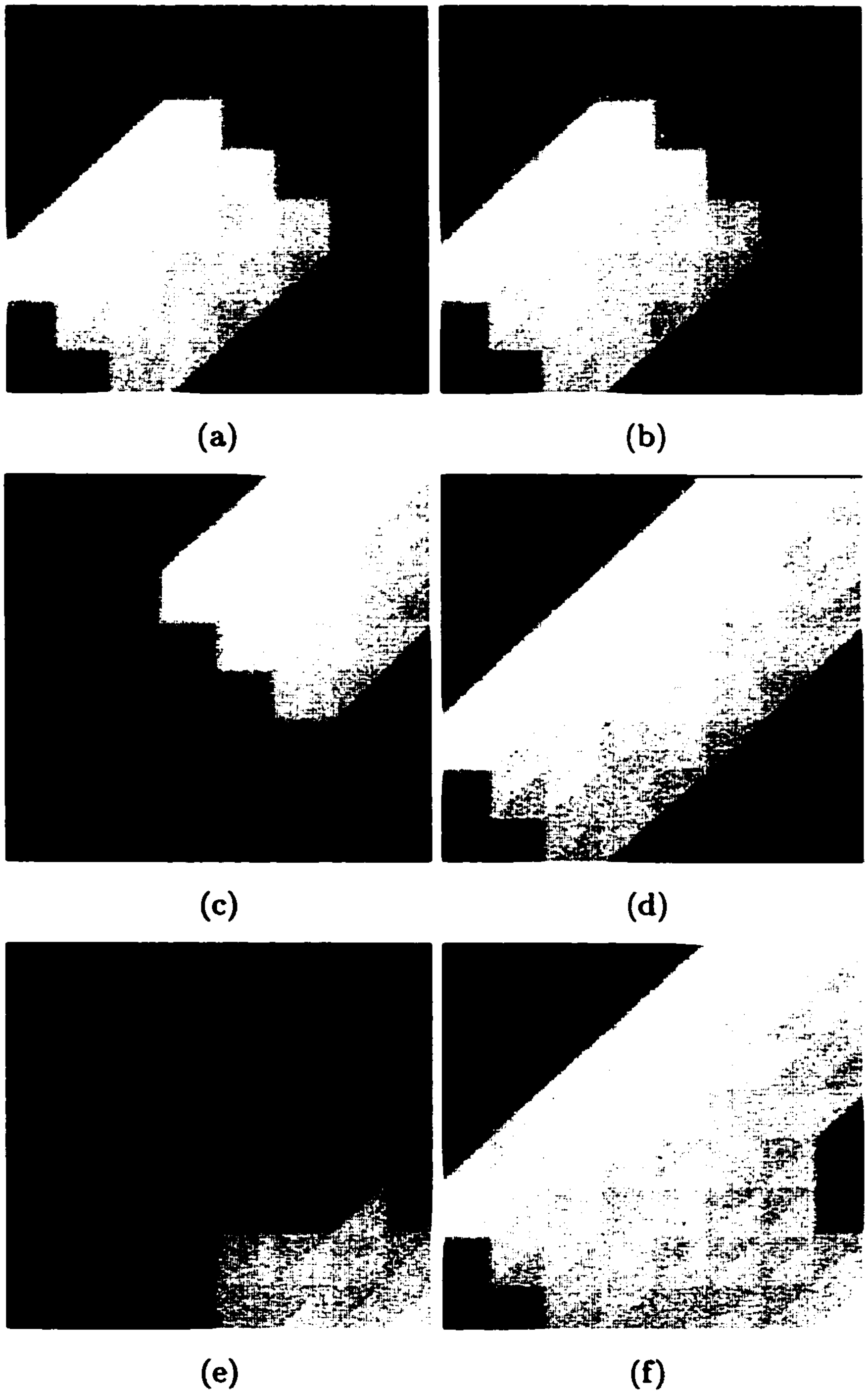
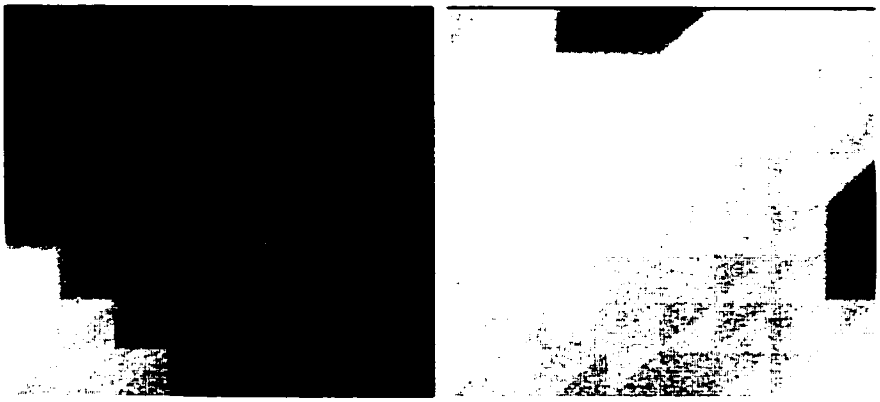


Figure 4.11: Convex Set Covering: 1-3

(a) (b)



(c) (d)

(e) (f)

Figure 4.12: Convex Set Covering: 4-6

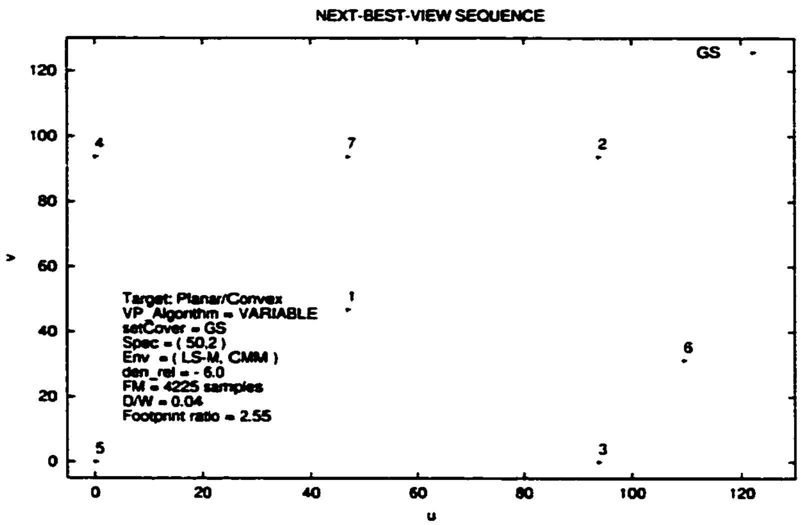


Figure 4.13: Next-Best-View Sequence (GS)

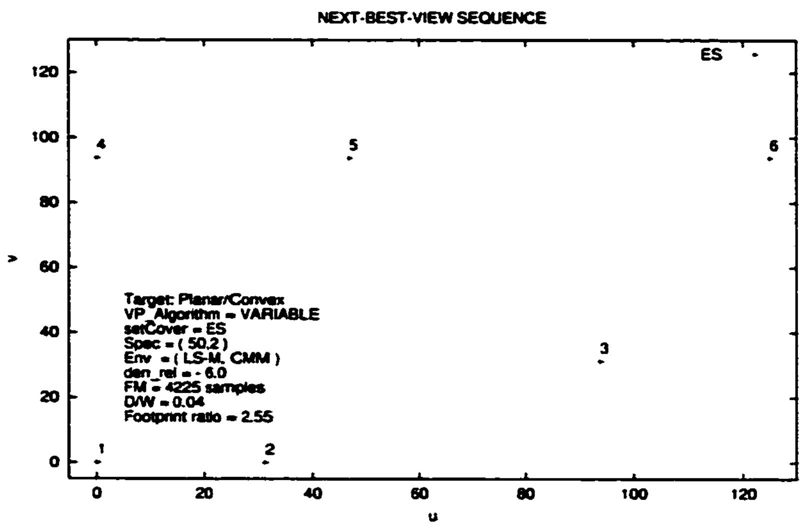


Figure 4.14: Next-Best-View Sequence (ES)

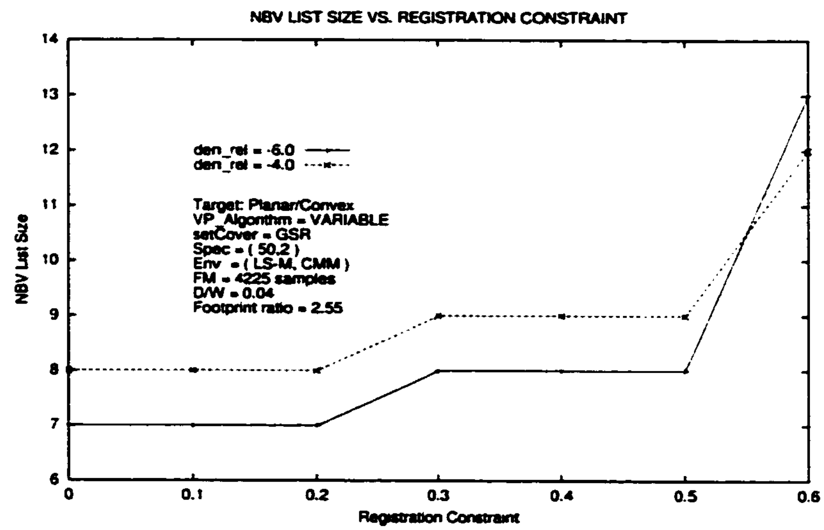


Figure 4.15: NBV List Size Versus Registration Constraint

4.4.3 Constrained Greedy Search

Next, we add an image-based registration constraint to the standard GS algorithm. The NBV is now defined as the view with the maximum FOM that also maintains connectivity of the registration graph. This is achieved by requiring the NBV to meet the correlation threshold (registration constraint) with at least one member of the current NBV set.

Figure 4.15 shows NBV list size variation with the registration constraint. The data confirms that the number of views necessary to cover the surface patch increases with constraint severity. While list size plateaus over some ranges of the constraint, the composition of the NBV set and the structure of the registration graph are evolving with the changing overlap requirement, as can be seen from Figure 4.16. The algorithm begins to fail beyond a certain point as the problem becomes over-constrained for the level of viewpoint space discretization. While this example of a moderately large and nearly planar patch illustrates the overlap-based approximation to the registration constraint, it is recognized that the viewpoint (image) sequence is not fully constrained for registration purposes in the absence of greater object shape complexity.

Figure 4.16 presents NBV positions in viewpoint space and their sequence, overlaid with the associated registration graph G_r for values of t_r in the range [0.1, 0.6] - that is, for an increasing overlap constraint. As the registration constraint is slowly tightened, we can observe subtle shifts in set covering and registration linkages. As specified, G_r maintains at least simple connectivity.

The matrix of raw correlation values and corresponding registration adjacency matrix A for the constrained set covering problem at Figure 4.16(b) for $t_r = 0.2$ are as follows, which can be compared to the unconstrained case.

$$\begin{bmatrix} - & 0.326 & 0.258 & 0.258 & 0.326 & 0.226 & 0.387 \\ 0.326 & - & 0.129 & 0.129 & 0 & 0.194 & 0.355 \\ 0.258 & 0.129 & - & 0 & 0 & 0.581 & 0 \\ 0.258 & 0.129 & 0 & - & 0 & 0 & 0.516 \\ 0.326 & 0 & 0 & 0 & - & 0 & 0 \\ 0.226 & 0.194 & 0.581 & 0 & 0 & - & 0.065 \\ 0.387 & 0.355 & 0 & 0.516 & 0 & 0.065 & - \end{bmatrix}$$

$$\begin{bmatrix} - & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & - & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & - & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & - & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & - & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & - & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & - \end{bmatrix}$$

To illustrate a set covering problem driven more by shape than relative size, Figure 4.17 presents the view plan generated for a deep cavity surface

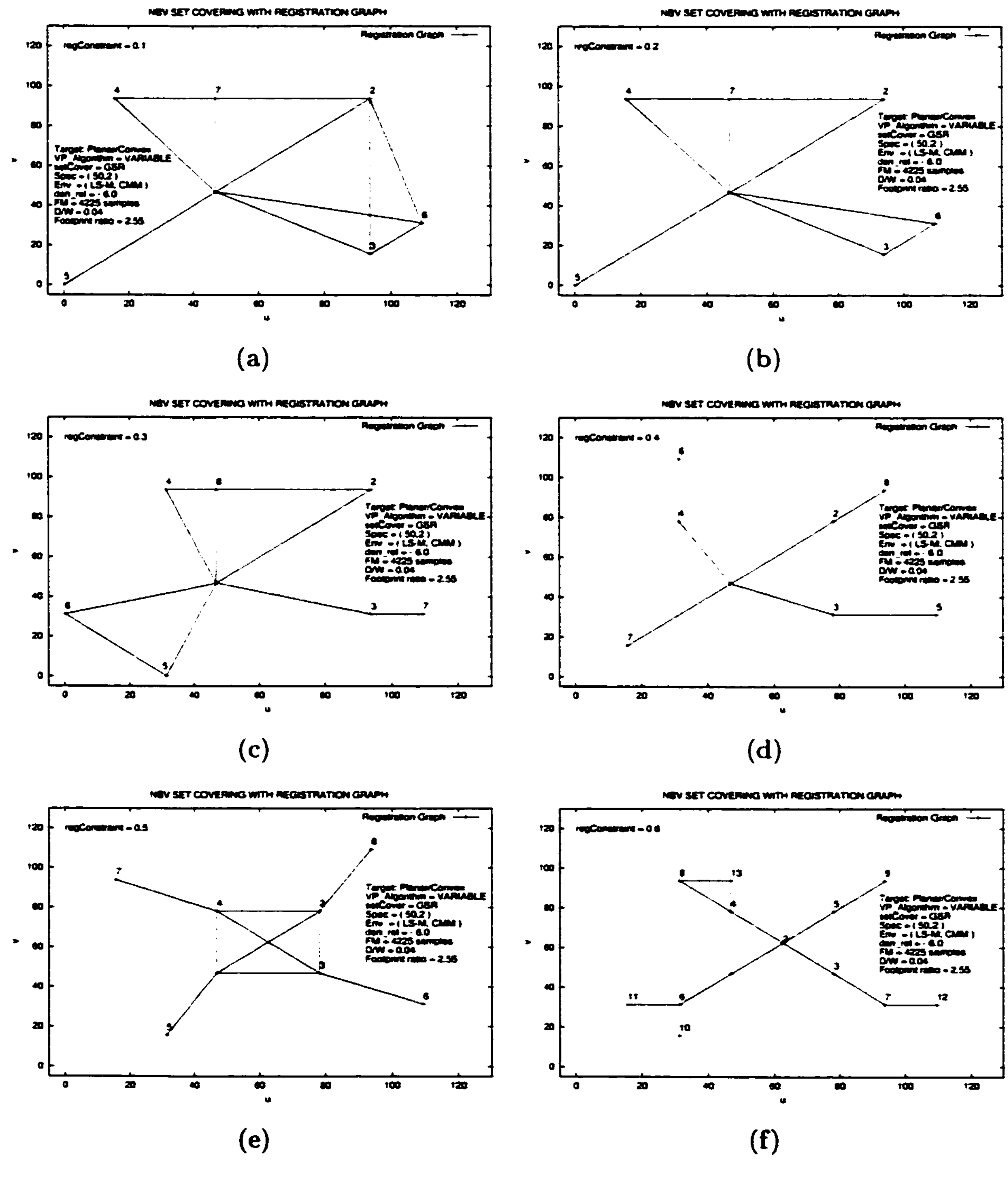


Figure 4.16: NBV Set Covering with Registration Graph

patch with $D/W = 0.5$, $r_f = 0.55$ and measurability matrix density = 0.637. The data is presented in extended Gaussian image (EGI) format[8] to highlight the orientation complexity of the patch. Two items are shown: the cavity mesh as represented by the angular distribution of surface normals and the axis component of each viewpoint in the NBV set overlaid with the registration graph. As expected, viewpoints are located on the opposite side of the Gaussian sphere from the mesh. To relate viewpoints to mesh orientations, their axis vector must be negated. Thus, viewpoints 2 and 3 (with the same axis but different rotation about that axis) are aligned with the cavity axis of symmetry while viewpoint 1 peers over the rim of the cavity into its depths. The viewpoint geometry is illustrated in Figure 4.18.

The registration graph is fully connected for $t_r = 0.2$ and the GS set covering is optimal ($m_v = 1.0$, $e_v = 1.0$). Figure 4.19 shows the cavity set covering sequence in the form of measurability projections. Again, individual shaded measurability projections are shown in the left column; composite measurability projections in the right. Black signifies "unmeasurable" while white indicates "measurable". Other shading is an artifact of the rendering program. In this case, unmeasurable regions are due to several factors - shadow effects or excessive incidence angles which either exceed threshold values or cause the estimated sampling density to drop below the specified level. This is an example of view planning influenced by a sensor model and a reconstruction specification.

[8] An EGI projection maps the orientation of 3D vectors (such as surface normals and the sensor boresight component of viewpoint orientation) to the corresponding point on a unit sphere. While a Mercator projection suffers from well known distortions in the polar regions, the projection provides a simple all-aspect presentation and the polar and Greenich axes can be selected to minimize distortion.

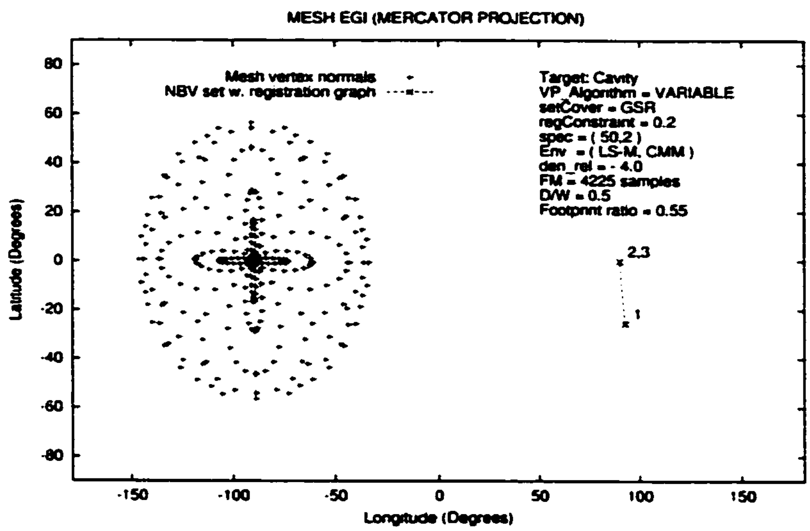


Figure 4.17: Mesh and NBV List EGI - Cavity

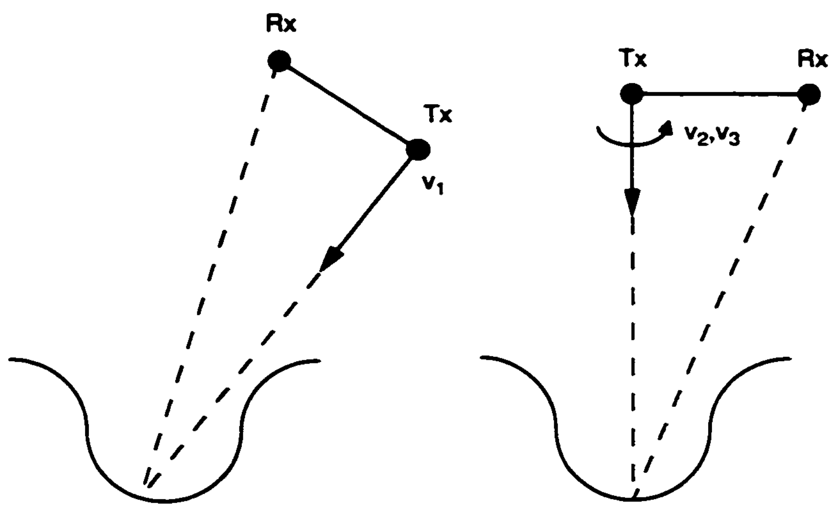


Figure 4.18: Cavity Next-Best-Views

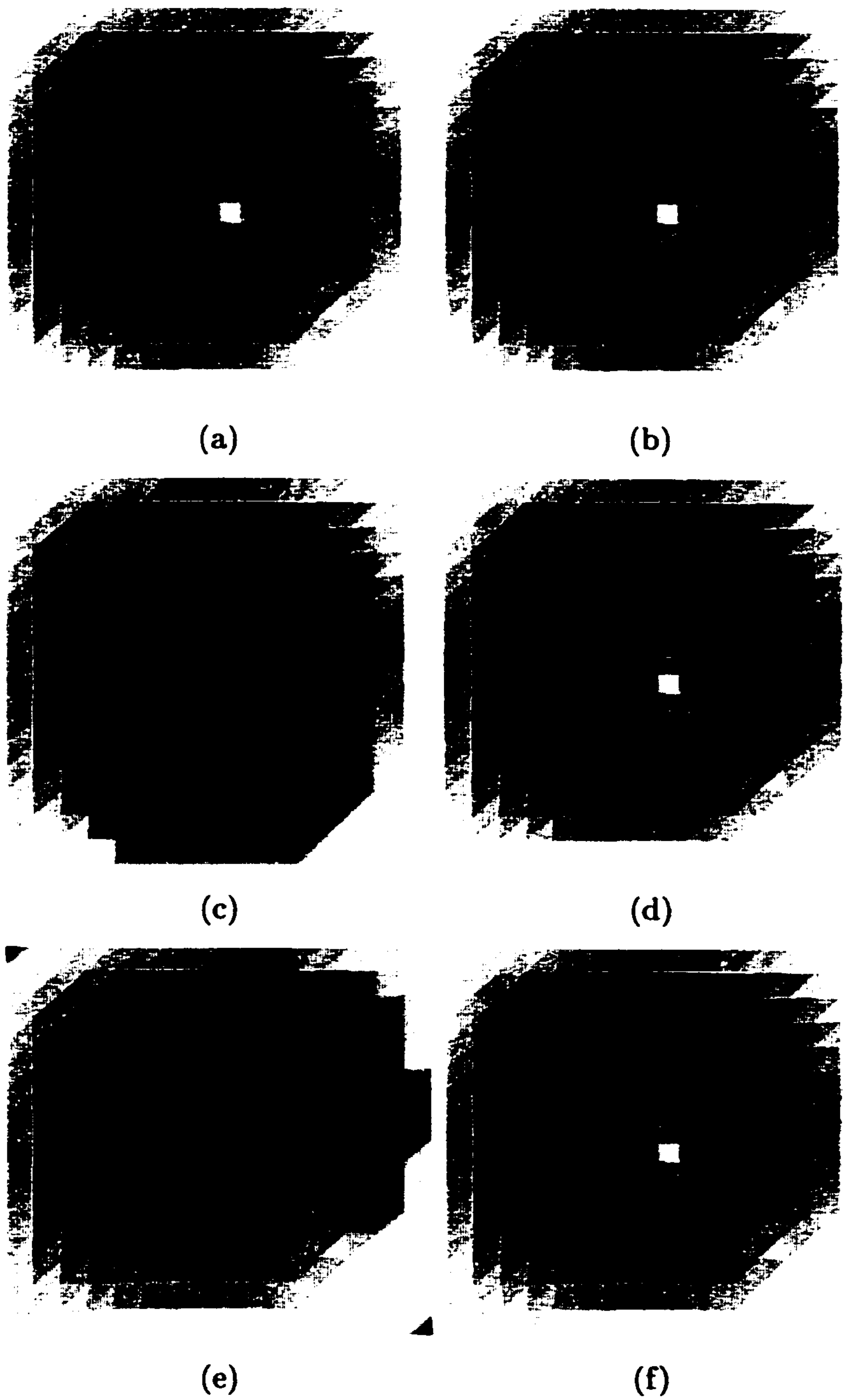


Figure 4.19: Deep Cavity Set Covering

4.5 Summary

Experiments reported here have revealed both strengths and weaknesses of set covering by greedy search. Two problems in particular are noteworthy. Firstly, close examination of Figure 4.10 reveals a small residual unmeasured region upon executing the complete NBV list. This artifact results from minor errors in surface normal estimation which translate into errors in viewpoint orientation and hence sensor footprint coverage. The problem manifests itself in small uncovered filets between verified measurability images. This can be corrected by slightly down-sizing the sensor frustum volume used for measurability analysis. Secondly, in the nearly planar patch example, we note a discrepancy between the minimum number of views determined by visual inspection (4) and the optimal result (6) achieved by exhaustive search of candidate viewpoints. This is because the candidate viewpoint set does not optimally span viewpoint space. While these results are still quite good for the current very low level of viewpoint space sampling (one viewpoint per rough model vertex), there is room for further optimization. In particular, useful trade-offs can be made between discretization levels in the orientation and position components of viewpoint space. These observations again illustrate that generation of a small set of high quality candidate viewpoints is the crux to efficiently solving the VPP. Viewpoint generation is addressed in Chapter 6.

The classic weakness of the greedy search algorithm, its irrevocable selection strategy [TG95], has minimal impact on many industrial and cultural object reconstruction applications. Frequently, the object size is comparable to or smaller than the sensor frustum ($r_f < 1$) and shape presents the dominant view planning challenge. Cases of large r_f are the exception rather than the rule. Consequently, measurability matrix density is often medium-to-high and solution set size is low relative to most set covering applications. In reconstruction, minor inefficiency resulting from a marginally longer NBV list is offset by fast and robust GS set covering computation plus added coverage insurance from redundancy. In contrast, inspection applications are generally prepared to trade longer off-line view plan computation for improved on-line view plan efficiency. In either case, we have presented a theoretical basis for computing optimal view plans by means of slightly more dense sampling of viewpoint space and more computationally intensive exact or heuristic set covering algorithms.

For object reconstruction, we find greedy search gives good set covering results and believe that computational resources are best allocated to quality viewpoint generation. Phrased differently, we believe the most difficult challenge of the view planning problem is framing the problem, not solving it. By "framing the problem" we mean determining an appropriate set of viewpoint variables and surface constraints and computing the associated measurability matrices. As we have shown, "solving the view planning problem" is analogous to solving the well-known set covering problem.

In conclusion, we have expressed a theoretical framework for the view planning problem as an integer programming problem including a registration constraint. The formulation is amenable to a variety of exact or approximate solution methods, depending on application requirements. We have also shown that good view planning results can be achieved by means of a simple and robust greedy search set covering algorithm modified to include a registration constraint.

Chapter 5

Sampling Object Surface Space

This chapter examines sparse sampling of object surface space with respect to sampling density, sampling noise and surface geometry.

5.1 Fine Modeling Process

Recall that our object reconstruction concept (Figure 3.3) involves a model-based, multi-step process of progressive refinement, separating exploration from precision measurement. Rapid, preprogrammed scene exploration produces a polygonal mesh *rough model*, an approximation of object geometry. Next, the rough model is decimated to a sampling resolution just adequate for view planning - the level being experimentally determined. Segmentation of the rough model may optionally be used for computational efficiency. The rough model is then used to develop a view plan for a subsequent stage of precision measurement in compliance with the model specification.

The current work concentrates on the fine modeling reconstruction stage shown in more detail at Figure 5.1. The process inputs are the rough model plus a model specification and imaging environment specification. The model specification defines reconstruction goals (measurement precision and sam-

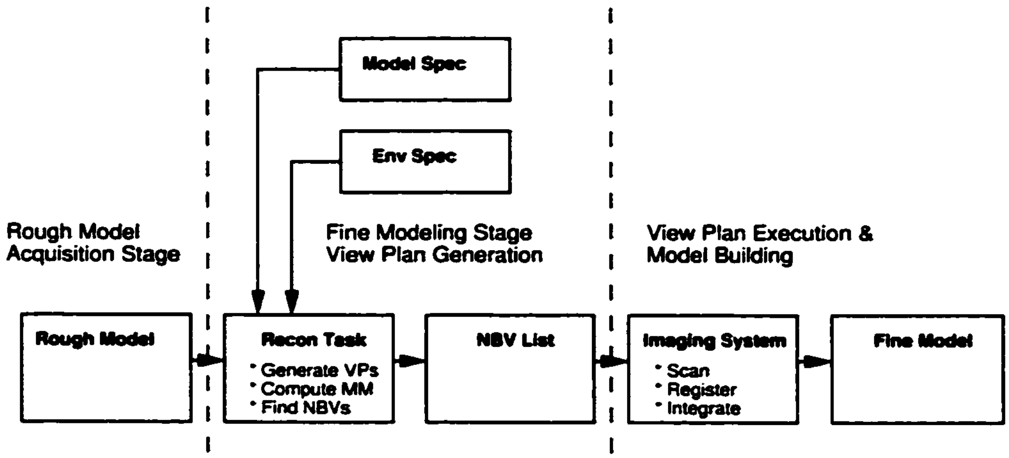


Figure 5.1: Fine Modeling Process

pling density) as well as configurable view planning parameters. The environment specification defines key sensor and positioning system parameters. The main view planning tasks at this stage are to generate candidate viewpoints, compute a measurability matrix and select the next-best-view list. That view plan[1] is sent to the physical imaging system to be executed.

For inspection, the rough model not only exists but is defined in detail. In object reconstruction on the other hand, we need to create a rough model quickly without laborious on-line procedures. The issue then, is -

How rough can the rough model be for effective view planning?

[1]The output of the set covering operation will be a list of viewpoints, generally ordered by imaging merit. Subsequent automation modules may chose to use that ordering, treat the NBV list as an unordered set or re-order the list by some other criteria, such as minimizing positioning system travel.

5.2 Experimental Process

The closed loop experimental procedure previously described at Section 4.4.1 was used with synthetic data to explore the impact of rough model decimation and sampling errors on view plan quality. A high density mesh such as the example at the left of Figure 4.8 serves as a surrogate fine model. The fine model is decimated to create coarse rough models at progressively lower sampling densities, after which surface noise is added. An example of a coarsely-sampled, noisy rough model is shown at the right in Figure 4.8. The rough model drives the view planning process, including viewpoint generation, measurability estimation and NBV list determination. The resulting view plans are executed against the parent fine model, closing the loop to provide verified measurability. The sensor model used in these experiments was previously described at Section 3.4. Measurability terms were defined in Section 3.3.2.

Experiments were conducted with synthetic cavity patches of various shapes as they represent a difficult view planning challenge. Synthetic shapes have the experimental advantage of being repeatable, quantifiable, modifiable and scalable. The principal input parameters are the size of the rough model relative to the objective fine model expressed in terms of relative sampling density, the level of rough model surface noise and the shape of the target surface patch. We examine verified measurability m_v achieved upon execution of the NBV list, cumulative measurability as a function of the number of views during the execution of the NBV list, size of the generated NBV list and measurability estimation error. Measurability estimation error (the difference between estimated and verified composite measurability) is an indicator of the reliability of the view planning process.

5.3 Geometry and Sampling Density

In this section we examine the impact of target shape and relative sampling density on view planning performance. Sampling density has received little attention in the view planning literature. Tarbox [TG95] observes that "The sampling points on the surface of the object should be sufficiently dense so that all surface regions, particularly the difficult-to-sense regions, are represented.". The associated sampling density analysis is restricted to grazing angle effects.

Shadow effects in surface concavities present a major scanning problem and view planning challenge when imaging with a bi-static sensor. The imaging geometry is determined by two key parameters: the optical baseline and the scanning range. The first is fixed while the second is variable. It is desirable to minimize the latter as geometric noise increases quadratically with range. Consequently, scans are normally taken close to the minimum range. The sensor modeled in these experiments has an optical baseline of 180 mm and a minimum scanning range of 142 mm - representative geometry for a mid-performance range camera.

Experiments were conducted with a variety of cavity shapes and rough model approximations. The optimal scanning zone viewpoint generation algorithm described in Chapter 3 was used. Consequently, in these experiments there is a 1:1 mapping between discretization levels in surface and viewpoint space. A greedy search set covering algorithm was used as described in Chapter 4. Figure 5.2 plots verified measurability as a function of rough model to fine model relative sampling density[2] ρ_{rel} for a cavity target over a range of geometry expressed in terms of depth-to-width ratio

[2]Relative sampling density is defined as $\log_2$ of the ratio between the sampling density of the rough and fine models.

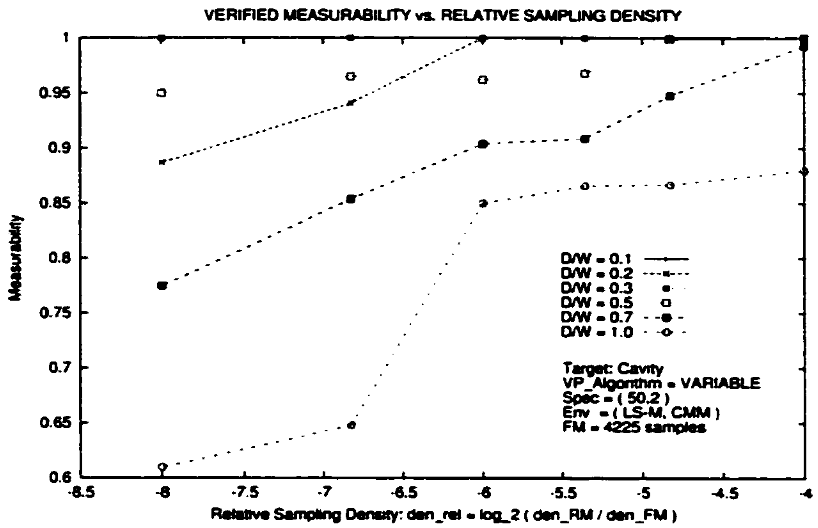


Figure 5.2: Verified Measurability vs. Relative Sampling Density

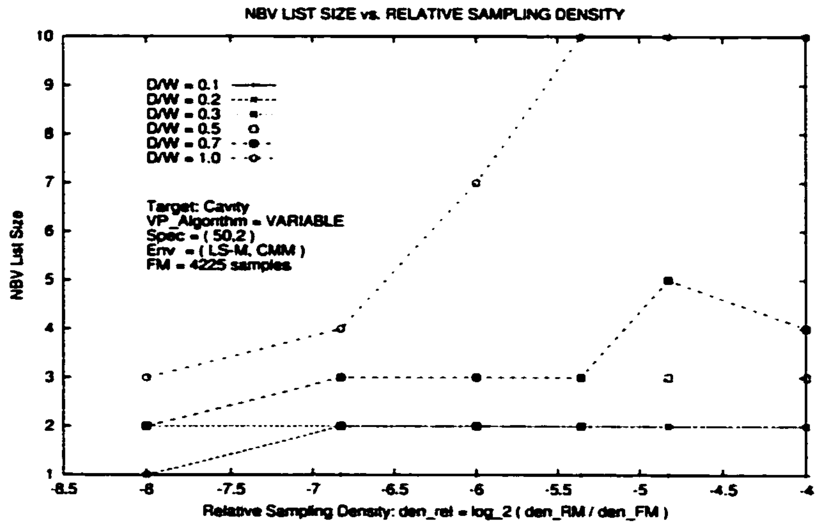


Figure 5.3: NBV List Size vs. Relative Sampling Density

D/W. In general, we would expect to encounter scanning and view planning difficulty for $D/W \geq 0.5$ for a sensor with a relatively long optical baseline such as the one modeled in these experiments. As expected, we observe that measurability improves with increasing ρ_{rel} as a consequence of a better candidate viewpoint set and higher fidelity rough model surface representation. Further, measurability is reduced with increasing geometric difficulty, although not always in a completely predictable fashion. Good measurability results are achieved up to $D/W \approx 0.7$ but drop off sharply for very deep cavities around $D/W \approx 1.0$. One anomaly we observe in this data set is an unexpected curve crossing at low relative sampling density (Figure 5.2). This phenomena, which is validated by close inspection of the data, is an indication of unreliable measurability prediction at very coarse rough model representations. Several inter-related factors are at work: errors in surface normal estimation impacting both the sampling of viewpoint space and surface grazing angle estimation, sparse sampling of both S and V, the spatial sampling phase and a set covering algorithm which does not presently distinguish between the robustness of next-best-views with identical figures of merit.

The second key performance indicator is the size of the NBV list. An approximate indicator of the number of views required for a patch is provided by the footprint ratio $\lceil r_f \rceil$, (Section 4.2.2). With $r_f = 0.55$, the cavity target used in these experiments could be measured by a single view at low D/W, which is confirmed by the experimental data. As the cavity departs from planarity, Figure 5.3 demonstrates that two views are soon necessary to capture the geometry. As few as three views are sufficient to measure all but the most extreme shapes. In these experiments, NBV list size was capped at ten views. The results show that, at very coarse relative sampling density levels,

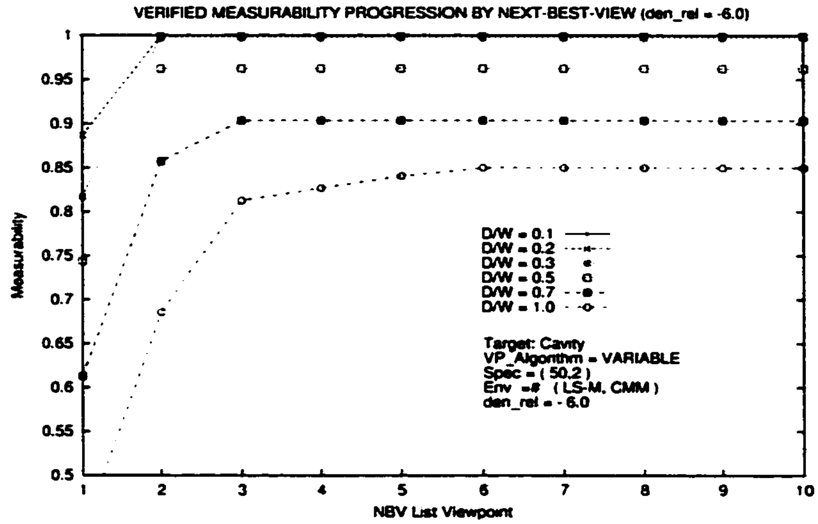


Figure 5.4: Measurability Progression with NBV Selection (ρ_{rel} = -6.0)

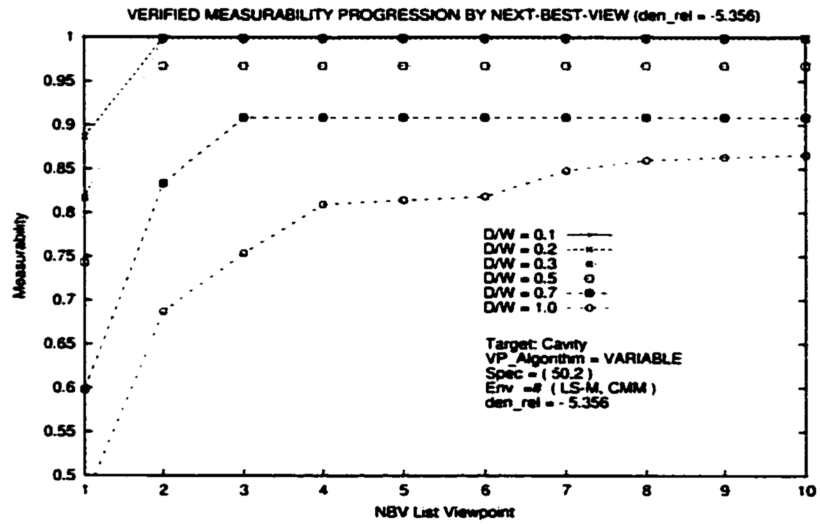


Figure 5.5: Measurability Progression with NBV Selection (ρ_{rel} = -5.356)

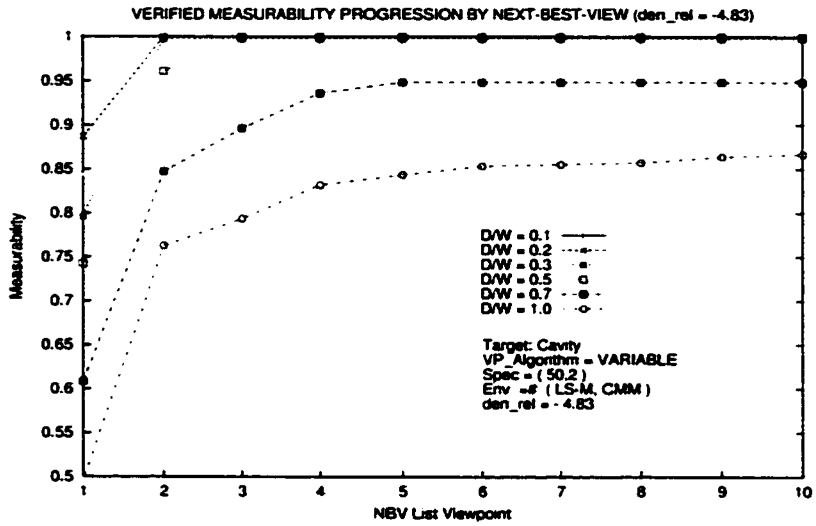


Figure 5.6: Measurability Progression with NBV Selection (ρ_{rel} = -4.83)

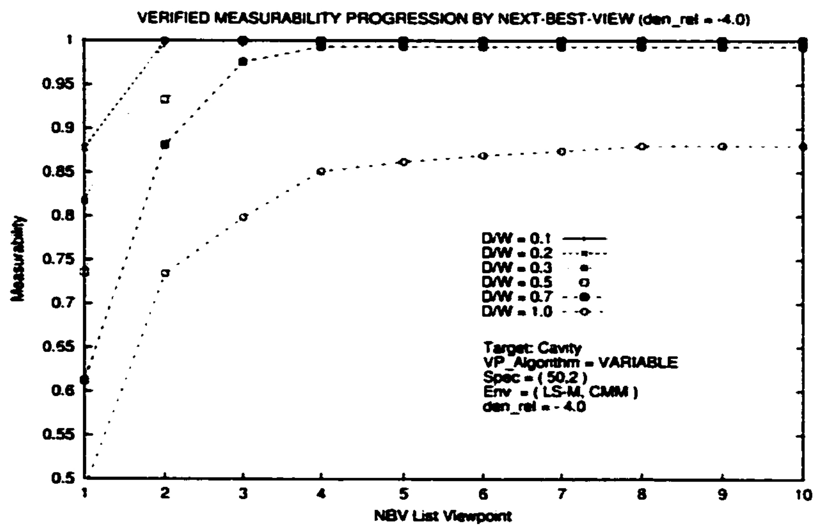


Figure 5.7: Measurability Progression with NBV Selection (ρ_{rel} = -4.0)

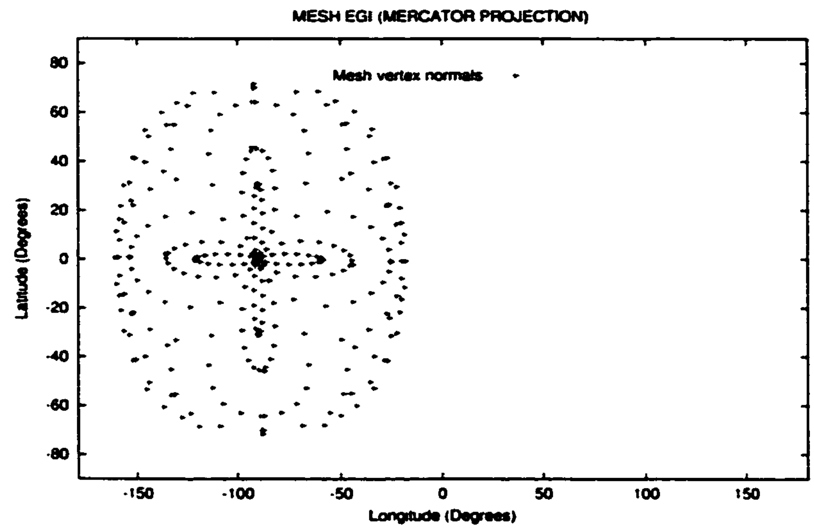


Figure 5.8: Mesh Extended Gaussian Image

the 3M algorithm underestimates the number of views required. However, beyond a shape-dependent sampling density threshold, generated view plans approach a stable length for near-optimal scanning.

Figures 5.4, 5.5, 5.6 and 5.7 show the progress of verified cumulative measurability with successive views from the NBV list for various rough model sampling densities. Viewpoints generated from the rough model are selected by a greedy search set covering algorithm on the basis of declining utility. Verified measurability progression is monotonic, usually but not always in decreasing increments due to minor errors in the prediction process. Further, almost all feasible measurement is achieved in a very small number of views, typically one to four for most shapes. For the deepest cavities, measurability gains beyond this number are marginal.

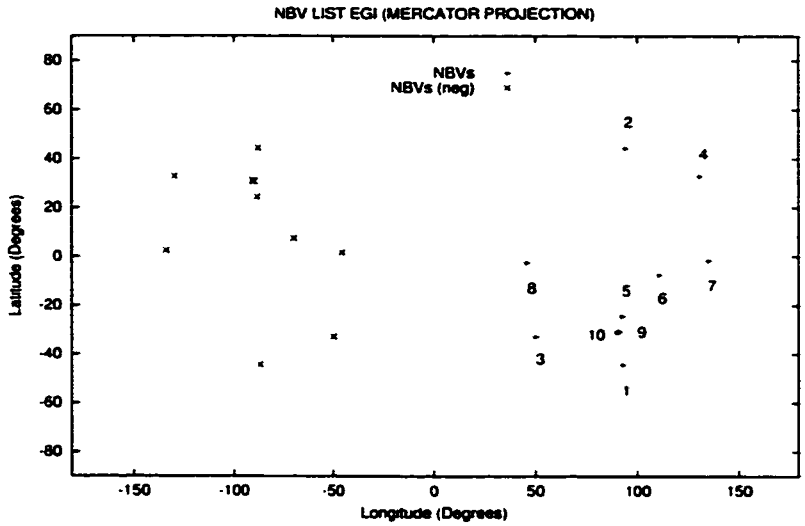


Figure 5.9: NBV List Extended Gaussian Image

An Extended Gaussian Image (EGI)[3] view of the distribution of surface normals for the very deep mesh ($D/W = 1.0$) used in the experiment at Figure 5.7 is shown at Figure 5.8 in Mercator projection format. In the particular test geometry examined here, we note that the mesh EGI exhibits an uneven orientation density. The surface normal orientation distribution for this shape falls into three distinct density zones: (1) a heavy concentration in the center corresponding to nearly flat and low curvature areas around the cavity rim and at its bottom, (2) another high density doughnut-shaped zone around the periphery of the EGI plot corresponding to low curvature areas on the cavity walls, and (3) an intervening low density zone corresponding to high curvature regions near the cavity lip and bottom.

Figure 5.9 presents an EGI view of the axis component of the NBV view-

[3]See the footnote to Section 4.4.3 for a description of the EGI projection.

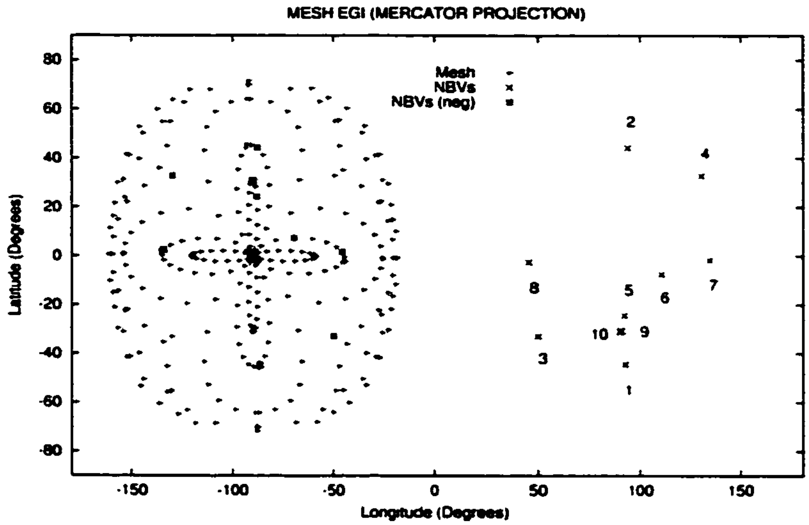


Figure 5.10: Mesh and NBV List EGI

point list generated for this target. As expected, NBV list viewpoints are situated on the opposite side of the EGI sphere to the target surface. The NBV list and its negative mirror image are displayed to permit correlation with the mesh. The NBV selection sequence illustrates how the algorithm maximizes measurement by first taking scans from the periphery in order to capture large surface areas near the rim. Measurement returns diminish with subsequent viewpoints clustering about the cavity depth axis as the algorithm attempts small gains in the furthest depths of the cavity.

In Figure 5.10, both the mesh and NBV EGI representations are overlaid, permitting correlation of the selected views with mesh orientations, albeit requiring a keen eye.

In summary, the relative sampling density level ρ_{rel} determines discretization intervals in both rough model surface space and viewpoint space. Increasing shape complexity demands a higher relative sampling density and vice-versa, but the relationship in 3D space is not as simple as the conventional sampling theorem for 1D communication signals. Our experiments show that a sampling level of $\rho_{rel} \approx -5$ (i.e. about 32 times lower than the fine model) provides good measurability prediction for concave shapes, the most difficult scanning challenge, for scanning geometries representative of contemporary range cameras. Lower ρ_{rel} is suitable for simpler shapes. For large nearly-planar surface patches, the sampling rate is constrained by coverage requirements related to the frustum cross-section, pose error considerations (dealt with in the next chapter) and registration/integration overlap requirements. For complex surface regions, the sampling rate is driven by visibility and surface normal estimation considerations. Finally, for typical surface patches and imaging geometries, most scanning coverage and measurement is achieved with a small number of views.

5.4 Rough Model Sampling Errors

By definition, the exploratory or rough modeling process is intended to execute quickly with limited a priori knowledge and without elaborate experimental setup. Therefore, we can expect rough model surface measurement error to be much higher than specified for the target fine model. Such errors are in addition to the intentional sparse sampling at this stage of the reconstruction process. This section addresses the impact of rough model sampling errors on view planning performance.

Geometric sensors prefer measurement along the local surface normal. The geometric noise component along the sensor boresight greatly exceeds transverse measurement error. Hence we select a simple model of rough model surface error based on a noise component along the local surface nor-

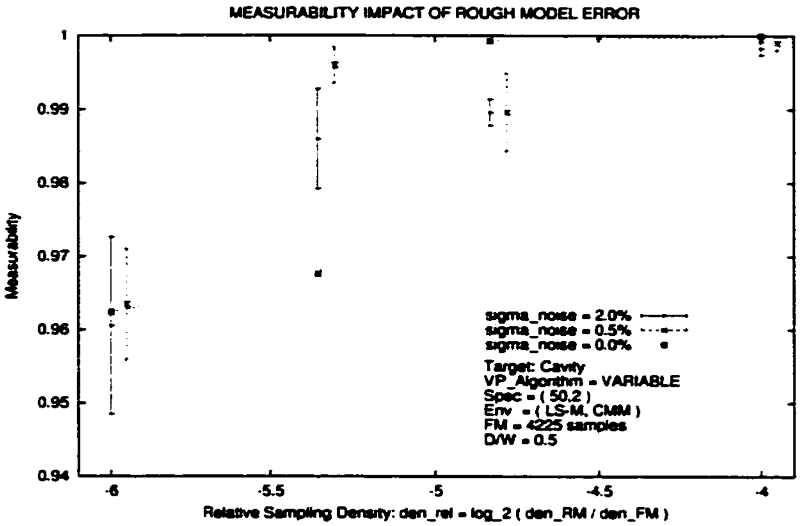


Figure 5.11: Measurability Impact of Rough Model Error

mal. The modeled noise is a zero mean Gaussian process with a standard deviation σ_{rm} expressed as a percentage of the sensor's minimum range. In the present work we do not address outliers, edge effects, corner reflections or sensor dynamic range limitations.

Two sets of experiments were conducted on a cavity test target of moderately difficult shape ($D/W = 0.5$) for a range of relative sampling densities using a random noise generator set at $\sigma_{rm} = 0.02R_{min}$ and $\sigma_{rm} = 0.005R_{min}$. The first noise level is quite high (i.e. $\approx 2.8mm$ for the sensor used in these experiments) such that it gives the rough model a noticeably crumbled shape (right hand figure at Figure 4.8). Figures 5.11 and 5.12 show the impact on verified measurability and the size of the NBV list of a noisy rough model in comparison to the noiseless case[4]. The noisy data is presented as average

[4]The data for $\sigma_{rm} = 0.005R_{min}$ is artificially separated slightly to the right for

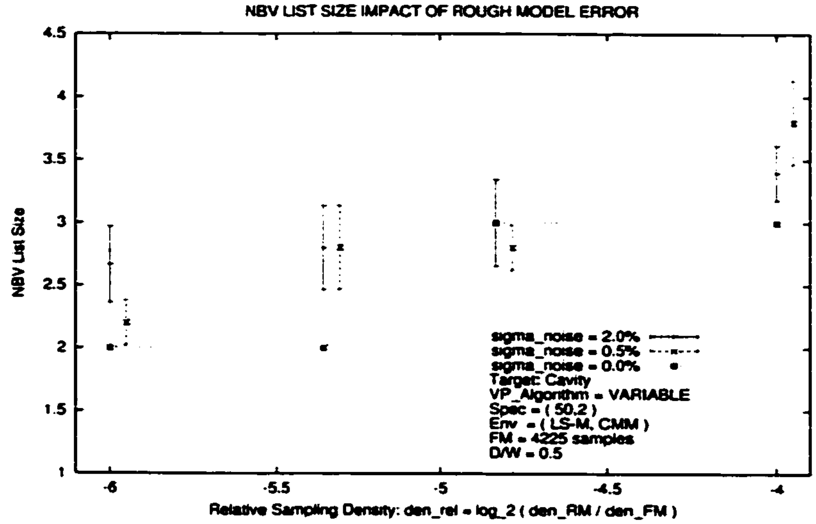


Figure 5.12: NBV List Size Impact of Rough Model Error

values plus or minus one sigma standard deviation.

Contrary to what might initially be expected, we find that rough model surface errors have a relatively minor impact on measurability. In some cases, the input errors actually result in an improved output. The reasons for this phenomena can be found in the impact on viewpoint generation. Random surface errors introduce a "dithering" effect on surface normal estimation. During the viewpoint generation process, this in turn randomizes the distribution of viewpoint positions as well as the axis component of viewpoint orientation. The twist component of viewpoint orientation (rotation about the camera boresight) is unchanged as it is set in reference to global shape characteristics. The net result of modest levels of rough model surface error is a randomized and possibly more uniform sampling of viewpoint space.

---

readability.

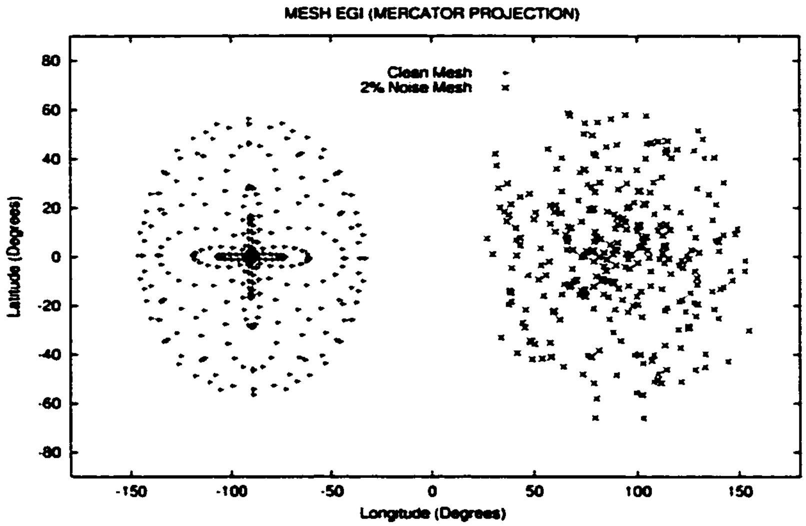


Figure 5.13: Clean and Noisy Mesh EGIs

The axis randomization effect is illustrated in Figure 5.13 which provides a side-by-side comparison of the EGI images of the cavity mesh used in the rough model surface error experiments at $\rho_{rel} = -4.0$. The clean cavity EGI is presented on the left while one instance of the same mesh perturbed by noise at $\sigma_{rm} = 0.02R_{min}$ is shown on the right. A similar presentation is contained at Figure 5.14 for $\sigma_{rm} = 0.005R_{min}$. The noisy cavity EGIs are noticeably more evenly distributed.

We can conclude that reasonable rough model random surface errors have the following effects:

- Measurability estimate uncertainty is slightly increased but average measurability is largely unchanged.
- On average, there is a slight increase in NBV list size. This additional scanning cost is offset by increased measurability robustness to other

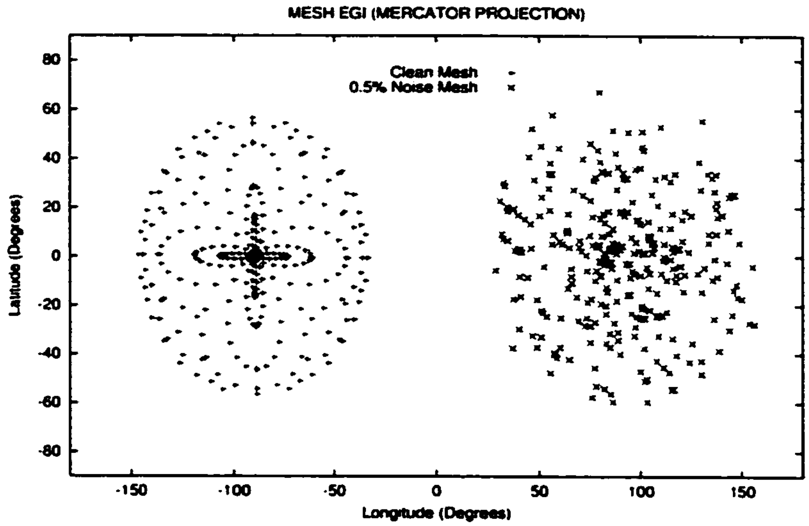


Figure 5.14: Clean and Noisy Mesh EGIs

error sources from a more extensive covering of viewpoint space.

5.5 Summary

It is instructive to examine actual versus estimated measurability performance. Figure 5.15 shows composite measurability estimated at the rough model viewpoint generation stage as a function of cavity shape and relative sampling density. At the rough model level, the estimation process is quite optimistic, predicting near perfect achievement of measurability goals for all but the most extreme cavities. Figure 5.16 presents the actual measurability achieved given the NBV list generated from the rough model. Notwithstanding the optimistic nature of measurability estimation at the rough model level, we note that the actual performance achieved is very good, particularly for $\rho_{rel} > -5$. The net measurability estimation error is presented in

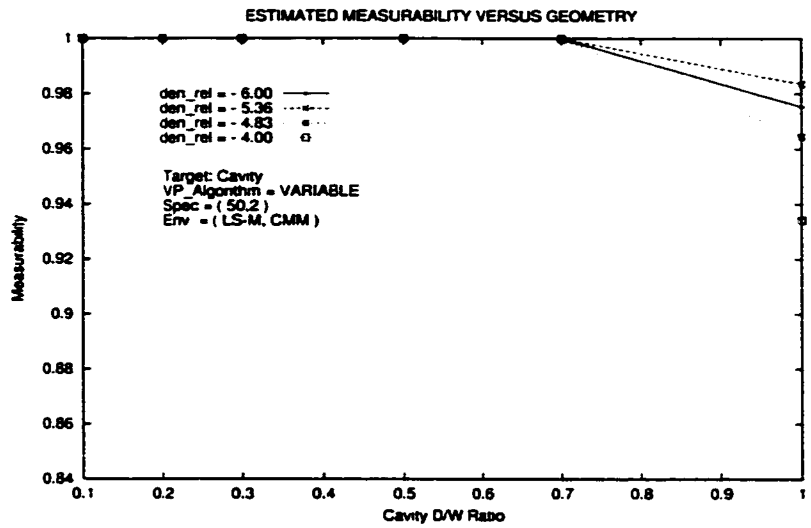


Figure 5.15: Estimated Measurability Vs. Cavity Geometry

Figure 5.17. The foregoing suggests that a rough model representation of about 250 triangles will provide good view planning results for fine model acquisition of a segment of about 8,000 triangles.

The complexity of the 3M algorithm is $\mathcal{O}(s^2v)$, where s and v are the sizes of the surface point and viewpoint sets, respectively. A key feature of the algorithm is sparse sampling in surface and viewpoint spaces. The size of the rough model is determined by segmentation of the target object into separately-targeted surface patches and by the degree of decimation feasible with each surface patch. The size of the rough model segment also influences the size of the corresponding viewpoint set. The experimental results of this section indicate that good results are achievable for a level of rough model decimation within the desired range. We have also shown that the algorithm is robust with respect to fairly high levels of rough model sampling error.

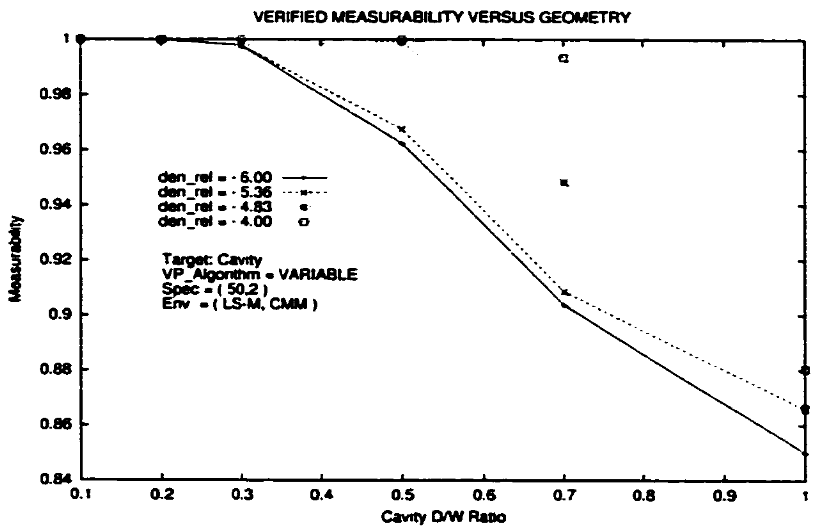


Figure 5.16: Verified Measurability Vs. Cavity Geometry

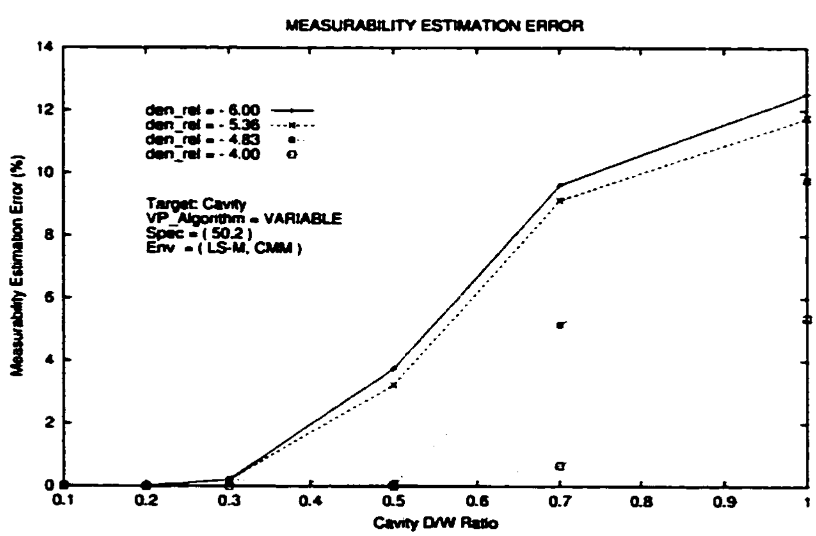


Figure 5.17: Measurability Estimation Error Vs. Cavity Geometry

Sparse sampling sharply reduces computational complexity resulting in sufficiently fast NBV computation to make automated reconstruction using this technique feasible. However, a sparse model also has disadvantages, notably the following:

- The quality of surface normal estimation deteriorates with decimation resulting in less optimum viewpoint pose (both position and orientation).
- Inaccurate surface representation with respect to viewpoints with high grazing angles near threshold leads to measurability estimation errors.
- Fewer candidate viewpoints are generated resulting in less dense sampling of viewpoint space. As the estimation process is probabilistic, the likelihood of sampling optimum viewpoint zones is reduced.

Synthetic targets were used for the experiments reported in this paper in order to control and measure the key variables. We are confident the sensor model employed at section 3.4 is a satisfactory basis for performance-oriented view planning for objects and imaging environments characterized by Lambertian scattering. For objects with shiny surfaces where specular reflections predominate, the limited dynamic range of the current generation of range camera electro-optical detectors will necessitate addition of a material reflectivity component to the rough model plus more detailed modeling of the laser source and detector.

Measurability estimates are random variables with complex statistics not easily characterized. Experimental view planning for a cavity segment, one of the most difficult shapes for scanning and view planning, indicates good results are achievable for a reasonable level of rough model decimation, surface errors within expected levels and sparse sampling of viewpoint space.

Chapter 6

Sampling Viewpoint Space

This chapter examines sparse and efficient sampling of viewpoint space.

6.1 Introduction

The goal of view planning is to derive an acceptably short list of viewpoints which collectively meet reconstruction objectives. There are several competing considerations. Computational complexity demands sparse sampling. Yet we need sufficient redundancy in the discretization scheme to provide adequate overlap for image registration and integration requirements, provide a buffer for pose error and not overly constrain the set covering operation that follows. Candidate viewpoints should have high prospective measurement quality with respect to sensor capabilities and the reconstruction model specification. Viewpoint space should be treated as the domain of generalized viewpoints, each consisting of sensor pose and a set of controllable sensor parameters. In the general case, we must be prepared for the high dimensionality of viewpoint space, $6D^+$, while catering to restrictions on degrees of freedom for specific imaging environments. Additionally, each degree of freedom of the positioning system and sensor will have an associated range of parameter values. As was made evident by some basic calculations in Section 3.6, brute force quantization of viewpoint variables is clearly unmanageable, particularly considering the fairly small size of view plan that suffices for most reconstructions. These issues raise an important question -

What is the optimal scheme for sampling viewpoint space?

We begin consideration of this question by examining conventional schemes for discretizing viewpoint space, then present a model-based approach - the "optimal scanning zone algorithm". This leads to discussion of viewpoint space sampling refinements and a second, improved viewpoint generation scheme - the "decoupled viewpoint generation algorithm".

6.2 Conventional Viewpoint Space Sampling Schemes

6.2.1 The Concept of a "Viewpoint"

Before proceeding further, some subtleties in the use and meaning of the term "viewpoint" merit elaboration. In this thesis, we consider a viewpoint to define a 3D frustum in world space. However, most triangulation range cameras (including the class of range camera modeled in the current work) are actually 2D sensors. They measure depth z over a 2D fan-shaped region swept by the laser in the sensor x-z plane. The third measurement dimension is achieved by physically moving the sensor perpendicular to the plane of light. This can be achieved by a sweeping motion along the camera y-axis (line-scan mode) or rotation about an axis parallel to the camera x-axis (cylindrical-scan mode). The geometry of a line scan range camera was illustrated at Figure 3.7. Range data gathered in such a manner will form a "range image" in the conventional sense - that is, a u-v grid of depth pixels.

A different data acquisition approach is to move the sensor through an arbitrary arc in space to collect a long swath of laser scan profiles. In general, the motion velocity vector need not have constant speed or direction. Sensor orientation may be constant or variable during the sweep. Lamb [LBG99] and Soucy [SCF98] employ this range data collection mode in contour following strategies using a CMM as the positioning system. In their setup, sensor position could be changed readily but orientation changes were costly in reconfiguration and recalibration time. Space carving techniques such as Papadopoulos-Orfanos and Schmitt [POS97] acquire range data in a similar manner.

Our use of the term "viewpoint" follows the first of these conventions. A viewpoint defines a single discrete position and orientation in space, with an associated constant 3D frustum shape. In addition to pose, a generalized viewpoint also has a set of configurable sensor parameters. We associate "range images" with viewpoints. Range image size (number of range pixels) is either fixed or is programmable over a narrow range of values.

A more appropriate term for the sensing operation commonly used with contour following and space carving techniques would be a "view swath". A view swath has a constant 2D frustum defined in the laser scanning plane. A view swath defines a continuous set of viewing positions and orientations over some arc in space, in effect a collection of 2D laser profiles. Neither the speed nor the direction of the velocity vector defining the sweep arc are constant, in general. View swathes are associated with either a set of 2D range profiles of arbitrary length or a 3D point cloud of arbitrary size.

The 3M view planning algorithm is designed to generate viewpoints. It could be extended to optionally generate view swathes for a contour following mode of operation, subject to some constraints, but that is out of scope of the present work.

6.2.2 Generalized Viewpoints

Most traditional view planning methods sharply constrain the dimensionality of viewpoint space in order to limit computational complexity. Few authors recognize the importance of sensor parameters in addition to pose. Tarabanis [TTA95] and Pito [Pit99] are exceptions to this observation. View planning for high performance sensors and difficult reconstruction tasks cannot ignore configurable sensor parameters.

6.2.3 View Sphere

The most common stratagem for constraining viewpoint space V uses a virtual, object-centered "view sphere" enclosing the object ([BA96], [Con85], [GVS98], [MF98], [MZH99], [RAS97b], [TG95], [YHS95]). This reduces the view planning problem (VPP) to the task of finding the best set of viewing directions. The viewing axis of any viewpoint on the sphere is oriented to the center of the frame of reference. Variable orientations about the viewing axis are rarely considered. This stratagem effectively reduces the dimensionality of V from $6D^+$ to 2D. A number of authors go further in constraining the problem, reducing V to a circular 1D domain, a simplification which tends to trivialize the problem and mask the underlying computational complexity issues. Many authors also appear to treat triangulation-based range cameras as mono-static devices, thus dismissing shadow-effects, an intrinsic sensor characteristic and key aspect of realistic view planning.

A few synthesis methods ([RAS97b], [TTA95]) treat the view sphere as a continuous domain. The majority, however, discretize the view sphere by a variety of techniques. This includes parallels and meridians [Con85], [ZMHN97], spherical distribution maps (SDMs) ([GVS98], [YT99], [YHS95]) or, the most popular, a tessellated icosahedron ([MF98], [MZH99], [SNSS92], [TG95], [TUWR97]), [ZMH98]. While it is well known that a uniform tessellation of a sphere does not exist, a tessellated icosahedron provides a close approximation. Morooka et al [MZH99] go one step further, transforming the geodesic tessellation to a flattened spherical array as an aid to improved efficiency in mapping orientation to geodesic cells.

Discretization by parallels and meridians offers simplicity at the expense of non-uniformity. A tessellated icosahedron provides an almost uniform sub-division but, depending on how it is used, may require complex computations regarding cell occupancy. The SDM approach offers intermediate performance - fast mapping of orientations to cells with reasonable cell uniformity, although inferior to the tessellated icosahedron.

Visibility analysis is an essential but computationally costly component of many view planning algorithms. Tarbox's approach [TG95] is unique in that viewpoint space is defined as the set of all points on the view sphere separated by a distance equal to the sensor optical baseline. For a tessellated icosahedron, this provides five or six orientations about the sensor boresight per viewing axis. In Tarbox's algorithm, view sphere vertices serve as locations for either the optical source or receiver. Ray tracing is conducted once per vertex rather than twice per viewpoint, thus reducing the cost of visibility analysis.

Almost all users of the view sphere strategy select a fixed stand-off distance such that the sensor frustum encloses the whole object. This may be necessary for some conventional surface-based techniques relying on occlusion edges to cue the next-best-view selection scheme. However, the strategy fails to take into account the important dependence of sensor performance on stand-off distance, a key feature of all triangulation-based range cameras. Also, it fails completely for imaging tasks for which object size exceeds frustum dimensions.

The principal advantages of the view sphere technique are two-fold: a reduction in the dimensionality of the pose component of viewpoint space from six to two, and relatively uniform sampling of the axis component of orientation (except for discretization by parallels and meridians). Its main disadvantages are as follows. To the degree that object shape is not spherical, viewpoint position is non-optimal. Measurement performance is further degraded if the view sphere radius must be set to view the entire object. Additionally, the view sphere ties orientation to position whereas there is more to achieving good viewing directions than merely selecting the orientation of the sensor boresight. The position from which the image is taken is also important. Further, the view sphere approach says nothing about the orientation of the camera around its boresight (the twist angle). For mid-to-high performance range cameras, the optical baseline typically subtends an angle in the range $[50° - 35°]$ at the sensor's optimum stand-off range. Consequently, the shadow effect, a phenomena integral to triangulation-based range cameras, is sensitive to rotation about the sensor boresight.

6.2.4 Intermediate Space

Pito [Pit99] separates visibility analysis of the object surface from that of the sensor by casting visibility images onto a virtual intermediate space representation, *Positional Space* (PS), positioned between the object's convex hull and the sensor. Visibility is verified by comparing the orientation of sensor ranging rays and object observation rays as encoded with respect to the local reference frame in Positional Space. However, PS representation is a lossy compression of pose information in which range-related information is lost. Stand-off range critically impacts sensor performance.

6.2.5 Summary

In summary, most view planning techniques constrain viewpoint space to two dimensions, a view sphere or view cylinder, and in some cases to a one dimensional viewing circle. Viewpoint positions are constrained to this surface or arc and the sensor boresight is fixed on the object center of gravity. Orientation about the sensor boresight is fixed (excepting the approach by Tarbox). The view sphere radius is also fixed, most commonly to a distance ensuring visibility of the complete object. The requirement to consider generalized viewpoints has not been widely recognized. The view sphere approach is mathematically attractive and provides reasonably uniform sampling in orientation but fails to optimize position nor to address important aspects of orientation optimization. Finally, the practicality of matching virtual spherical viewpoint domains with the capabilities and limitations of real positioning systems, real sensors and real objects is rarely considered.

6.3 Model-based Viewpoint Generation

6.3.1 Optimal Viewpoint Zone

A cross-section through the frustum of a range camera is illustrated in Figure 6.1. The angular field of view in the scanning plane generally lies in the $[20^\circ - 50^\circ]$ range, with 30° being representative. The best performance (measurement precision and sampling density) is achieved along the camera boresight at the minimum scanning distance R_{min}. To the first order (Chapter 3, Equations 3.3 and 3.4), depth measurement precision deteriorates quadratically with range and inversely in proportion to the cosine of the grazing angle while lateral measurement precision deteriorates linearly with range.

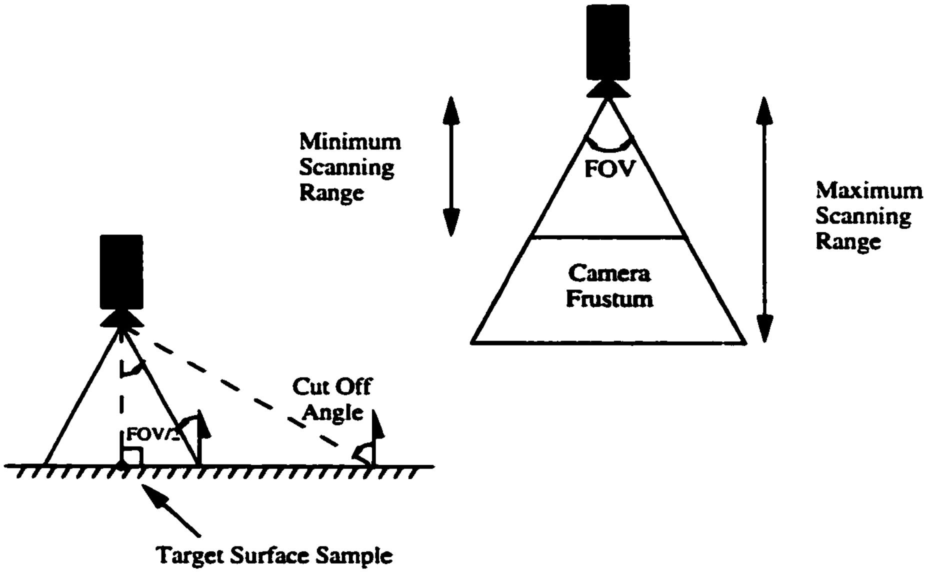


Figure 6.1: Optimal Stand-off Viewpoint

Optimal measurement precision will therefore be achieved with the camera positioned at the minimum scanning range perpendicular to a targeted surface point and oriented in opposition to the local surface normal. This configuration ensures that both incidence and scanning angles are zero. Optimal rotation of the camera about the boresight will be influenced by local shadow effect considerations. The set of all such optimal viewpoints corresponding to the point-by-point dilation of object surface points along their local surface normals forms an optimal scanning surface for the parent object.

Consequently, we are led to the concept of an *optimal viewpoint zone* $\mathbf{S_o}$ for an object as illustrated with a 2D cross section in Figure 6.2. While it is convenient to depict the optimal viewpoint zone as a surface, in practice it designates a narrow zone with somewhat fuzzy boundaries. This is a consequence of the non-uniform variation of measurement precision and

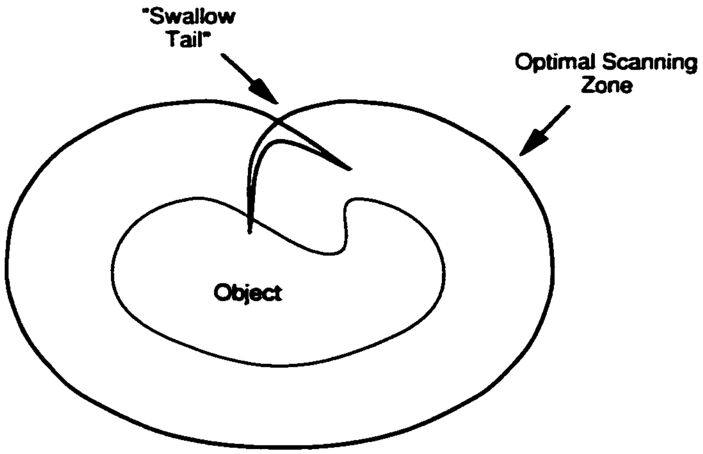


Figure 6.2: Optimal Viewpoint Zone

sampling density within an image as well as trade-offs between stand-off distance, lateral displacement and grazing angle. The optimal viewpoint zone $\mathbf{S_o}$ is smooth and well behaved in regions corresponding to planar and convex portions of the object surface. However, depending on object surface curvature and stand-off distance, "swallow tail" effects[1] may occur in regions corresponding to surface concavities. In these areas, $\mathbf{S_o}$ may violate a collision avoidance buffer zone or even intersect the object surface or be occluded by a portion of the object. $\mathbf{S_o}$ undergoes non-linear and non-uniform changes (including stretching and compression as well as possible tangent and curvature reversals) and is generally not manifold. However, the local topology of the surface remains unchanged by dilation.

[1] These effects are similar to those occurring with evolving wavefronts. See, for example, the treatment of evolving interfaces by level set methods at [Set96].

6.3.2 Optimal Scanning Zone Algorithm

The concept of an optimal viewpoint zone leads us to propose the following near-optimal viewpoint generation scheme. As the application is high performance object reconstruction, we assume a positioning system with six degrees of freedom and an adequate range of motion. Mobility constraints will be addressed later. An ideal candidate generalized viewpoint $(\mathbf{v}, \lambda_s)$ is created for each surface point - specifically for each vertex of the rough model mesh. It is convenient to consider pose orientation as the concatenation of two rotations - boresight axis and rotation about the boresight axis (twist). Thus, for view planning purposes, we use pose $\mathbf{v}$ = (position $\vec{p}$, axis $\vec{a}$, twist t_w). In our software implementation, homogeneous transforms are used internally where appropriate.

Each rough model vertex is dilated along the local surface normal by a stand-off distance $f_d R_o$, where f_d is a stand-off adjustment factor and R_o is the optimum scanning range, usually equal to the sensor's minimum range. f_d is an experimentally-derived constant slightly greater than one to allow for errors in measurability estimation and sensor positioning. The dilated surface vertex becomes the position of the corresponding candidate viewpoint. The viewpoint's axis is set to the negative of the dilation vector.

Setting the sensor twist orientation is less straight forward as it is influenced by shadow effects arising from object shape over a broader surface region. As an initial approach, we attempt to set the twist angle to minimize local shadow effects. With segmented rough models, one approach is to characterize the shape of each surface patch and align the sensor baseline with the patch's major axis. In the unsegmented case, the twist angle is random.

Viewpoints violating the collision avoidance zone are filtered out with the expectation that associated surface regions will be covered by adjacent near-optimum viewpoints on $\mathbf{S_o}$. Pose constraints arising from positioning system mobility limitations are applied on a case-by-case basis. Essentially, the optimum viewpoint will be mapped to the nearest feasible pose. The mapping strategy must be tailored to the specific set of mobility constraints unique to each imaging environment. Finally, configurable sensor parameters are optimized for each candidate viewpoint, based on the model specification, system capabilities and characterization of the rough model surface. In the case of the line-scan range camera studied in the present work, only one sensor parameter is configurable - scan length.

We postulate that the set of viewpoints thus generated provides a near-optimal sampling of viewpoint space for the purpose of solving the view planning problem[2]. Viewpoint parameters may be optimized on a per-viewpoint basis or on a per-patch basis. This discretization scheme reduces the dimensionality of viewpoint space from $6D^+$ to 2D. View planning optimization is conducted over the domain of generalized viewpoints. With a $1:1$ correspondence between surface points and candidate viewpoints, the complexity of the optimal scanning zone algorithm for viewpoint generation is $\mathcal{O}(s^3)$.

Figure 6.3 summarizes the algorithm's salient features. The rough model is decimated to a sparse level of surface sampling as discussed in Chapter 5. One optimal viewpoint is generated per rough model mesh vertex. Consequently, there is a $1:1$ mapping between surface point set S and viewpoint set V. Thus, there is an equivalent sparse level of discretization in the position component of viewpoint space, with a single orientation per viewpoint.

[2]Given our sensor model, individual viewpoints generated in this manner are optimal with respect to the targeted surface point for the case of an object surface with Lambertian scattering properties, which is the case considered in the present work. In the case of specular surfaces, the imaging geometry is more sensitive and must take the location of both the optical transmitter and receiver into account.

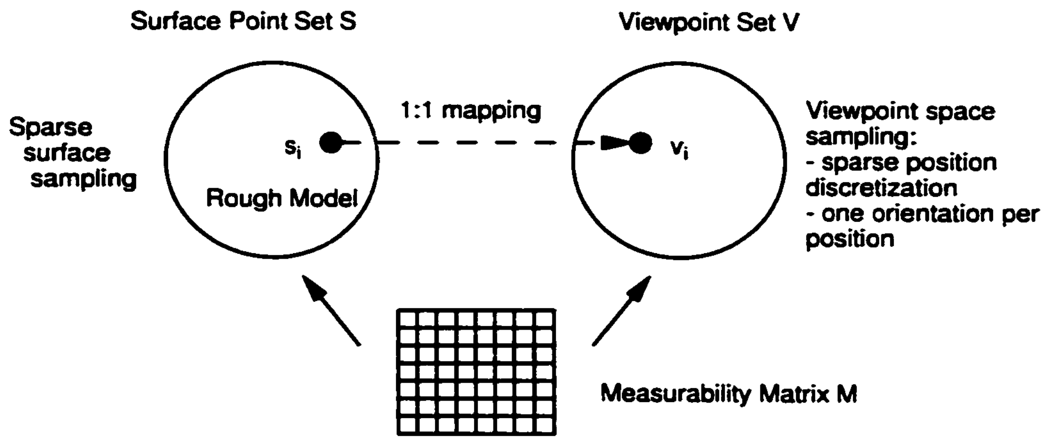


Figure 6.3: Viewpoint Generation - Optimal Scanning Zone Algorithm

Once pose is determined, sensor parameters are optimized on a per-viewpoint or per-segment basis. The algorithm samples a 2D subset of the $6D^+$ viewpoint space.

6.3.3 Surface Segmentation

This section and the following two sections describe enhancements to the basic optimal scanning zone view planning algorithm.

The portion of the surface of a 3D object covered by a given viewpoint is limited by the size of the sensor cross-sectional footprint and by occlusion. Clearly, large areas on the back side of the object are not visible. Visibility analysis is the critical measurability matrix computation. Yet much analysis is wasted on surface regions out of view. Additionally, some shapes (such as cavities and holes) present more difficult imaging and view planning challenges than less geometrically complex regions. Consequently, an effective complexity-reduction strategy is to manually or automatically segment the object surface into patches, preferably along the boundaries of major geometric features. A measurability matrix is then computed and the set covering algorithm is run separately for each segment. The final view plan is the concatenation of the next-best-view lists for each segment. Segmentation amounts to a partitioning of the overall measurability matrix.

Assuming uniform segmentation into c segments, the complexity of computing one measurability matrix is $\mathcal{O}(s^2v/c^3)$. Then, the overall complexity of computing the set of measurability matrices is $\mathcal{O}(s^2v/c^2)$. If we follow the $1:1$ mapping of surface points to candidate viewpoints in accordance with the strategy at section 6.3.2, these expressions are $\mathcal{O}(s^3/c^3)$ and $\mathcal{O}(s^3/c^2)$, respectively.

This substantial reduction in complexity is gained at a price, however. Visibility analysis is conducted locally with respect to the geometry of one segment at a time. Poor segmentation may result in selection of views which are, in fact, occluded from surface regions otherwise deemed measurable. Additionally, the overall view plan is bound to be somewhat less efficient due to redundant coverage of segment edges. The shape of segment boundaries also influences set covering efficiency. Nevertheless, computational savings from segmentation are substantial. In the inspection application case, segmentation will often be appropriate as many inspection tasks are interested in only a portion of the object surface.

6.3.4 Single Best View

If segmentation is used, a profitable sub-strategy is to delay generating a full slate of candidate viewpoints in favour of first attempting to find a single viewpoint or small set of viewpoints potentially able to scan the entire segment. We generate this candidate "single best view" set as follows.

- *Axis* - Set to the reverse of the patch's average normal.
- *Position* - Dilate the patch center vertex along the patch's average normal such that the patch falls completely within the sensor FOV.
- *Twist* - Coarsely sample twist angle with d_t quantization levels. Intervals of 90° may be adequate at this stage.
- *Parameters* - Match configurable sensor parameters, such as scan length, to the patch geometry as observable from the candidate pose(s).

Measurability matrix columns are computed and the SCP algorithm is run on this initial candidate viewpoint set. If successful, generation and test of the full slate of more finely sampled candidate viewpoints is aborted. If unsuccessful, viewpoint generation and measurability matrix computation continues using the normal optimal scanning zone algorithm.

6.3.5 Optimal Scanning Zone Algorithm Variations

In this section, we consider two versions of the optimal scanning zone algorithm with subtle variations in the details of viewpoint generation.

In the current work, two classes of range camera were modeled - the line-scan configuration as previously described (x-axis scanned electronically, y-axis scanned mechanically by means of sensor motion) and the random-access configuration (both x- and y-axes scanned electronically from the same point in space). In our implementation, the following sensor parameters were optionally reconfigurable: field of view in the x-z plane Φ_x (the laser scanning plane), field of view in the y-z plane Φ_y (applicable to random-access cameras), scan length L_y (applicable to line-scan cameras) and image size (N_x-by-N_y range samples).

Two variants of the optimal scanning zone algorithm were implemented. Named the CONSTANT and VARIABLE algorithms, these differed in how the following variables were set: stand-off distance R, twist angle t_w and sensor parameters λ_s.

CONSTANT Algorithm The stand-off distance is set to $R = R_o f_d$ as previously described. The rotation angle t_w of the sensor about the viewing axis is set such that the sensor optical baseline (which defines the sensor y-axis) is aligned with the plane containing the viewpoint axis and the major axis of the targeted surface patch. If the camera is reconfigurable, parameters applicable to the modeled sensor class are optimized. As presently implemented for the reconfigurable case, this means image size is maximized and the applicable combination of field of view and scan length variables defining frustum shape for that sensor class is maximized consistent with conservative tests to satisfy the specified sampling density. Variables are subject to minimum and maximum limits contained in the imaging environment specification. If not reconfigurable, sensor parameters are set to nominal values which are also defined in the environment specification.

The defining characteristic of the CONSTANT algorithm is that stand-off distance and sensor parameter values are optimized once for each segmented surface patch. Viewpoint pose (position, axis, twist angle) is optimized on a per-viewpoint basis.

VARIABLE Algorithm The VARIABLE algorithm proceeds as above with the following differences. First, sensor parameters are optimized on a per-viewpoint basis. Secondly, sensor twist angle is set to maximize the distance of the receiver position from a plane defined for the surface patch (Figure 6.4(a)). This enhances visibility of most surface features. Thirdly,

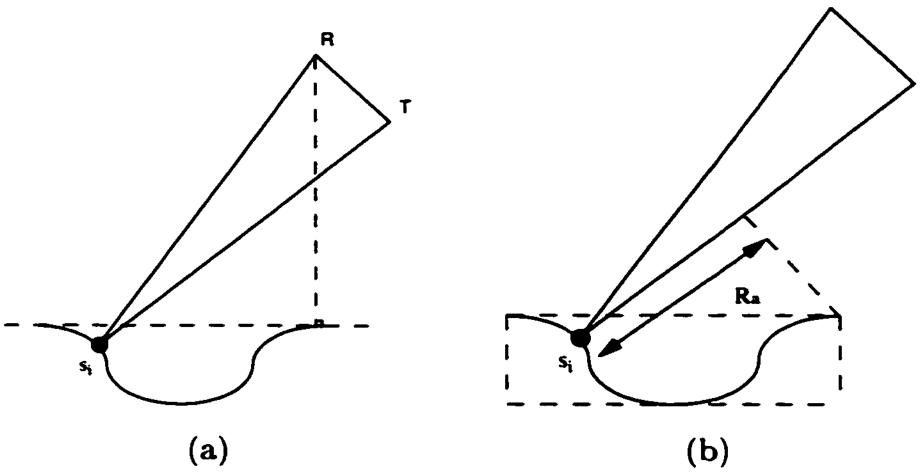


Figure 6.4: VARIABLE Algorithm

stand-off distance is adjusted on a per-viewpoint basis with an allowance to observe portions of the surface patch nearer to the observer than the targeted surface point. The rationale for this adjustment is to avoid unnecessarily clipping surface regions falling into the sensor field of view due to the sensor's minimum range. Then, $R = (R_o + R_a)f_d$, where the stand-off adjustment range R_a is the range difference along the sensor boresight between the targeted surface point and the projection onto the viewing axis of the corner of the segment bounding box nearest to the sensor (Figure 6.4(b)).

Performance Comparison All experimental data presented in the thesis to this point have used one of the preceding two variants of the optimal scanning zone algorithm for viewpoint generation.

Experiments showed that the two variants had similar net performance in terms of view plan quality and efficiency, although the VARIABLE version of the algorithm had a slight overall advantage. Computational efficiency differences between the variants are of only minor concern as the crux remains computation of the measurability matrix following viewpoint generation. Both versions generate the same number of candidate viewpoints.

While both versions provide good performance, a common area for improvement is setting the twist component of viewpoint orientation. Shadow effects are particularly sensitive to the sensor orientation about the viewing axis i.e. the twist angle t_w. With the CONSTANT algorithm in particular, we found we were unable to automatically characterize the major axis of a surface feature sufficiently accurately and reliably for optimal view plan efficiency.

6.3.6 Computational Efficiency

The Tsimshian stone mask (Figure 6.5)[3] is an interesting object to illustrate the computational efficiency of model-based view planning with a complex physical shape. The mask is carved from stone in the form of a thick shell. A high quality mesh model of the mask acquired by the NRC colour range camera was decimated to a lower resolution mesh (402 vertices, 800 triangles) and then partitioned into front and back segments. We selected the front mesh (292 vertices, 527 triangles) as an example rough model. Contrary to what one might conclude from a casual inspection of the segment, much of the fine-detail shape structure is located along the rim of the mask rather than in the face itself.

[3]The Tsimshian stone mask is regarded as one of the masterpieces of northwest coast art in the collection of the Canadian Museum of Civilization (VII-C-329). It was collected at the Tsimshian village of Kitkatla on the northwest coast in 1879 by I.W. Powell.

Figure 6.5: Tsimshian Stone Mask

An Extended Gaussian Image (EGI)[4] view of the front segment (Figure 6.6) shows mesh vertex normals to be widely distributed in direction, even though the segment pertains mainly to the front of the mask. For this example, the imaging environment was specified as a line-scan range camera with an error-free positioning system. The range camera was modeled on a specific commercial sensor ($\Phi_x = 27°, b = 180mm, R_{min} = 142mm, R_{max} = 407mm, L_{y(Max)} = 300mm, N_x = N_y = 256$). The reconstruction specification called for a precision of 50 μm and sampling density of 2 s/mm^2.

[4]See the footnote to Section 4.4.3 for a description of the EGI projection.

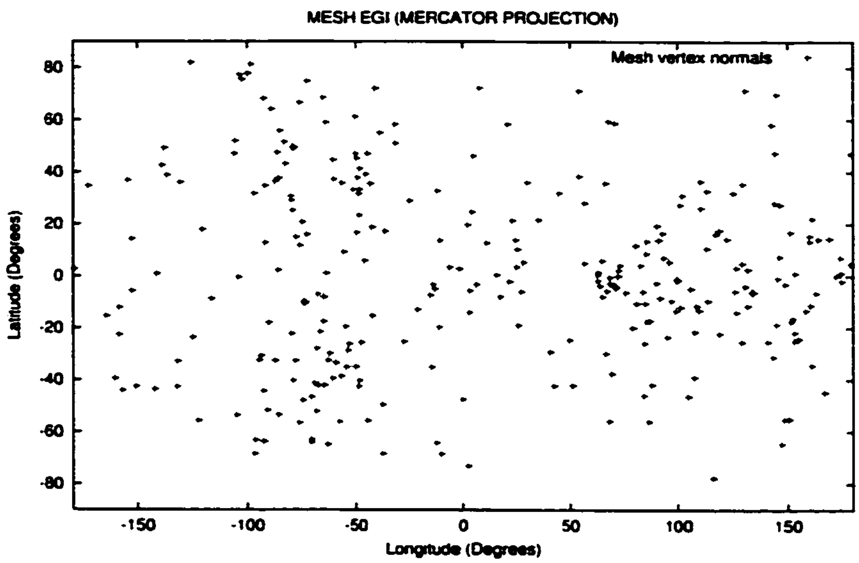


Figure 6.6: Tsimshian Stone Mask Mesh EGI

While the broad shape characteristics of the object are clear, it is not immediately obvious to a human operator what sequence of views would efficiently cover the object completely, yet alone fully satisfy the performance-oriented reconstruction objective.

Viewpoints were generated with the VARIABLE algorithm. Using a constrained greedy search set covering algorithm with a registration constraint of $t_r = 0.1$, the 3M view planning algorithm produced a view plan completely measuring the segment with 13 views. The view plan is shown in the EGI projection of the negative of viewpoint axes at Figure 6.7 to allow direct correlation with the mesh. The four left-most viewpoints in the EGI plot at Figure 6.7 cover the center of the face and the two cheeks - a gently curved region of large surface area but modest vertex density. The remaining seven viewpoints show the algorithm concentrating on the more geometrically com-

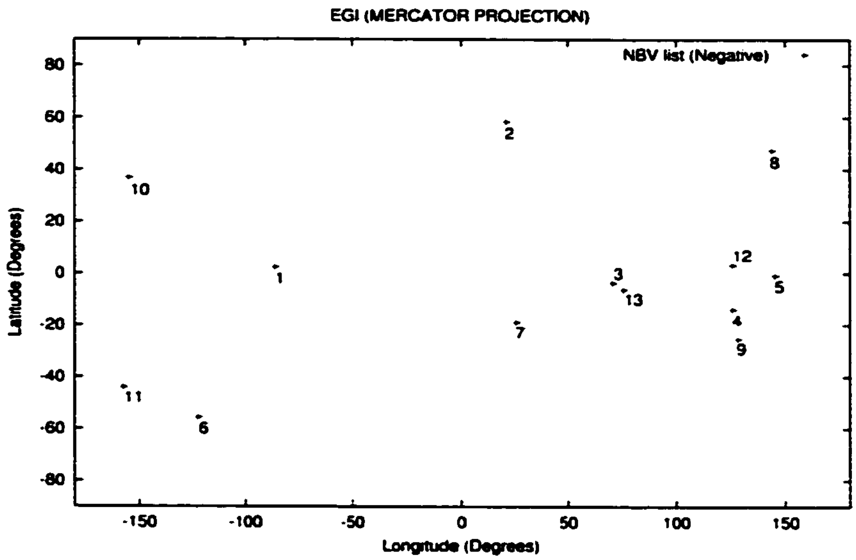


Figure 6.7: Tsimshian Stone Mask NBV List (Negative) EGI

plex regions behind the ears, along the rim of the mask, the forehead and underneath the chin.

Computing the measurability matrix took 63.2 minutes on a 150 MHz SGI O2 IP32 processor with 128 MBytes of main memory. The greedy search set covering operation required approximately one second. The size of the measurability matrix was 292 surface points by 296 viewpoints[5].

In a separate experiment, we show the verified set covering of the view plan computed for the rear segment of the mask at Figures 6.8 and 6.9. The view plan for the rear segment consists of five views. As before, individual shaded measurability projections are shown on the left while cumulative measurability projections are shown on the right. The view plan achieves

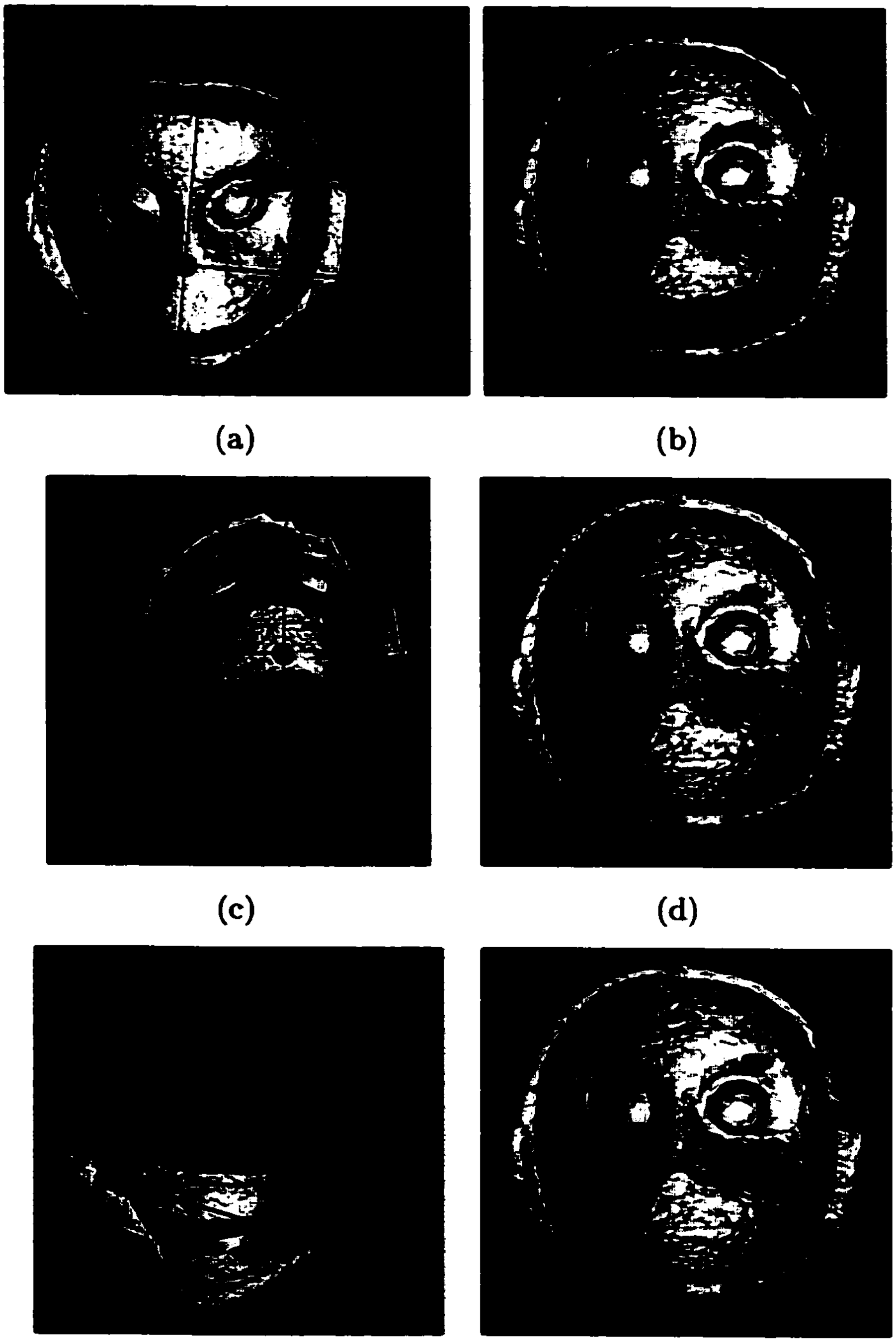
[5]The additional four viewpoints were the "single best view" set.

a verified measurability of $m_v = 0.97$, measuring all surface points within specification excepting a few spots on the steep side walls of the mask cavity.

The 3M algorithm was coded in C++ with a structure and coding style intended for a robust and flexible research platform. A conservative estimate of speed-up possible through code optimization would be a factor of two. CPU speeds continue to advance rapidly. At the time of writing, 1.5 GHz CPUs are commonly available at consumer prices. Current parallel computing technology allows construction of powerful low cost clusters of 20 or more CPUs. Experiments in our laboratory [LGG01] on parallelization of the related registration problem have shown linear speed-up for a cluster of 40 processors. Thus, a conservative estimate of the potential speed-up with currently available, low cost computing and networking technology is a factor of 2 x 10 x 20 = 400. Such a configuration would produce a view plan for the example problem in about 10 seconds or, alternatively, create a view plan for a mesh segment of size $s \approx 2100$ within our target time limit of one hour. These figures show that our algorithm will achieve the stated computational efficiency goals with currently available, low cost parallel computing technology.

6.3.7 Summary

The algorithm guarantees that the position and viewing axis of each candidate viewpoint is optimum for its targeted surface point. Additionally, the viewpoint is nearly optimum for an area around the targeted surface point and will provide specification-compliant measurements over some broader surface region. Setting of the viewpoint twist angle is reasonably good but could be improved. Improvements will require either improved rough model surface characterization (difficult) or more fine-grained sampling of the twist

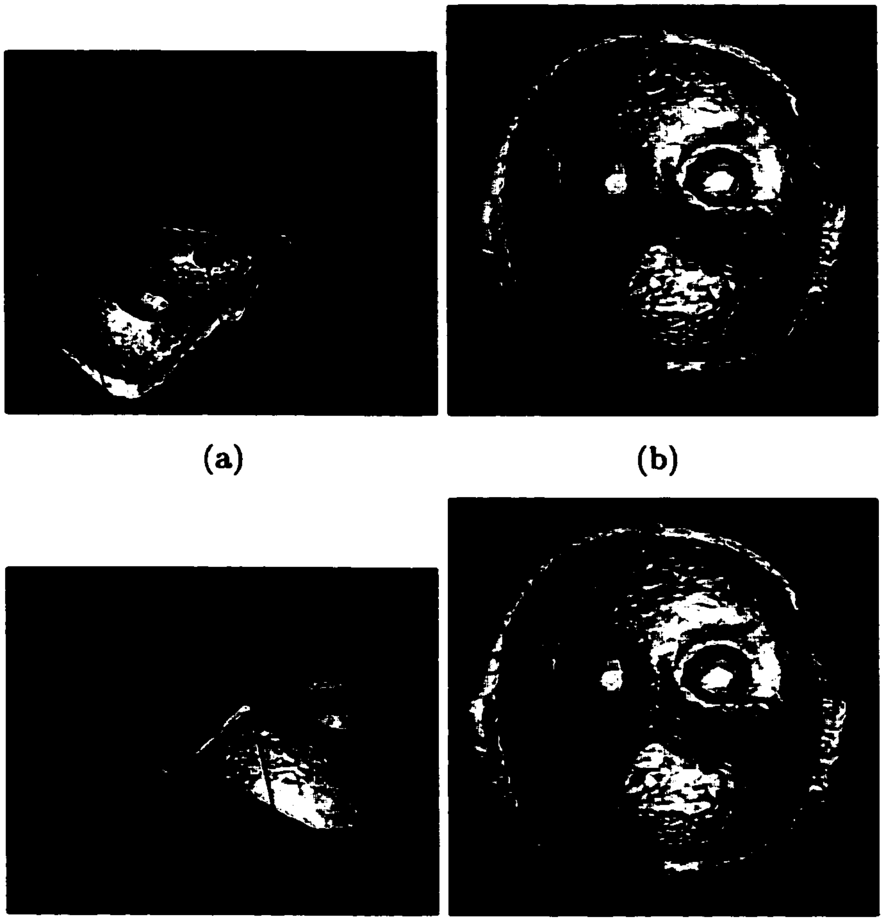


(a) (b)

(c) (d)

(e) (f)

Figure 6.8: Tsimshian Stone Mask Set Covering - Back (1-3)



(a) (b)

(c) (d)

Figure 6.9: Tsimshian Stone Mask Set Covering - Back (4-5)

rotation angle.

One counter-argument to the proposed strategy is that there are benefits in some circumstances to placing the sensor in the "far field" region. Specifically, the triangulation angle between receive and transmit rays decreases with increasing stand-off distance. In theory, this improves the sensor's ability to observe the interior of deep cavities. In practice, however, improved visibility is gained at the expense of a larger stand-off distance and sometimes more acute incidence angles over the regions of greatest interest, both of which degrade measurability. With a judicious choice of stand-off adjustment factor f_d, we conclude that the proposed viewpoint generation scheme is near-optimal. Nevertheless, it is amenable to further refinement.

In a somewhat similar approach, Prieto [PRBL99] synthesizes optimum viewpoints for an inspection application. Surface normals are first computed from NURB surfaces in the CAD model. If not occluded from the targeted surface patch, the viewpoint is set along the surface normal at the optimum stand-off distance, else it is moved until visibility is achieved at a minimal grazing angle. It is unclear from the work whether the bi-static nature of the range camera and therefore shadow effects are considered.

The contour following approach of Lamb [LBG99] implements a set of semi-automated contour following modes of operation in which sensor pose is subject to a number of constraints. These include maintaining a stand-off distance such that the surface lies within a high accuracy zone in the center of the sensor depth of field (DOF), scanning perpendicularly to the surface and minimizing the number of orientation changes. In general, these are mutually exclusive constraints. Swathes are limited in extent to regions over which the stand-off and grazing angle constraints are met within some tolerance. Soucy's contour following approach [SCF98] is somewhat similar. Both schemes generate view swathes rather than discrete viewpoints.

The distinguishing characteristic of our viewpoint generation approach is that it is performance-oriented - that is, based on sensor performance models and quantified measurement performance objectives.

6.4 Viewpoint Space Discretization Strategy Refinement

This section examines refinements to the strategy for discretizing viewpoint space. We begin by developing a simplified model for viewpoint correlation and then examine trade-offs in sampling the position and orientation components of viewpoint pose. With this background, we then present a new viewpoint generation algorithm - the DECOUPLED algorithm.

6.4.1 Simplified Viewpoint Correlation Model

Viewpoint correlation was defined in Chapter 4 at Equation 4.1. We also examined viewpoint correlation for some simple object shapes in Section 4.2.3. We observed that neighbouring viewpoints are highly correlated and that correlation falls off with displacement and is modified by surface shape.

For efficient viewpoint generation, we need some rules of thumb for spatial sampling rates in viewpoint space. With this purpose in mind, it is useful to develop a simple viewpoint correlation model with respect to translation in the optimal scanning zone. We show the geometry of such a model at Figure 6.10. Here, we consider an optimized frustum shape which is square in cross-section. We also simplify the model to image a planar target and consider frustum overlap area as the major factor in viewpoint correlation. We define the sensor cross-section f_s as $f_s = \sqrt{F_s}$ where F_s is the sensor cross-section area or footprint at the optimal stand-off distance (Section 4.2.2). If

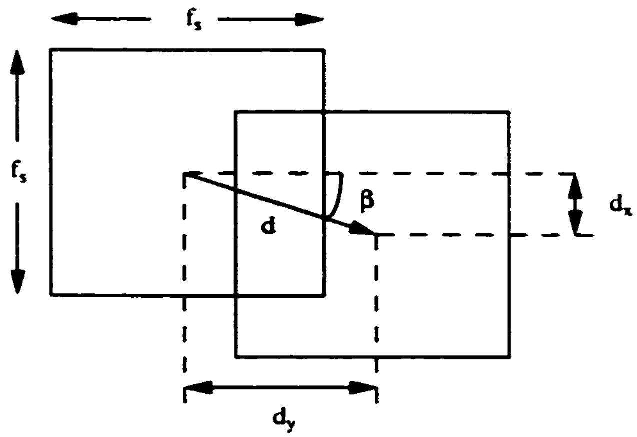


Figure 6.10: Simplified Viewpoint Correlation Model

a viewpoint is off-set a distance d at a random angle β, then the relative overlap area $A_{rel} = A/{f_s}^2$ is

$$\begin{aligned} A_{rel} &= \frac{(f_s - |d_x|)(f_s - |d_y|)}{{f_s}^2} \\ &= 1 - \frac{d}{f_s}(|C\beta| + |S\beta|) + \frac{d^2}{{f_s}^2}|C\beta S\beta| \qquad (6.1) \end{aligned}$$

for $d < f_s$. As β is a uniform random variable over $[\pi, -\pi]$, $E[|C\beta|] = E[|S\beta|] = 2/\pi$ and $E[|C\beta S\beta|] = 1/\pi$. Then, the average relative overlap area is

$$\mu_{A_{rel}} = 1 - \frac{4}{\pi}\frac{d}{f_s} + \frac{1}{\pi}(\frac{d}{f_s})^2. \qquad (6.2)$$

Figure 6.11 shows we can simplify the relationship further by the approximation

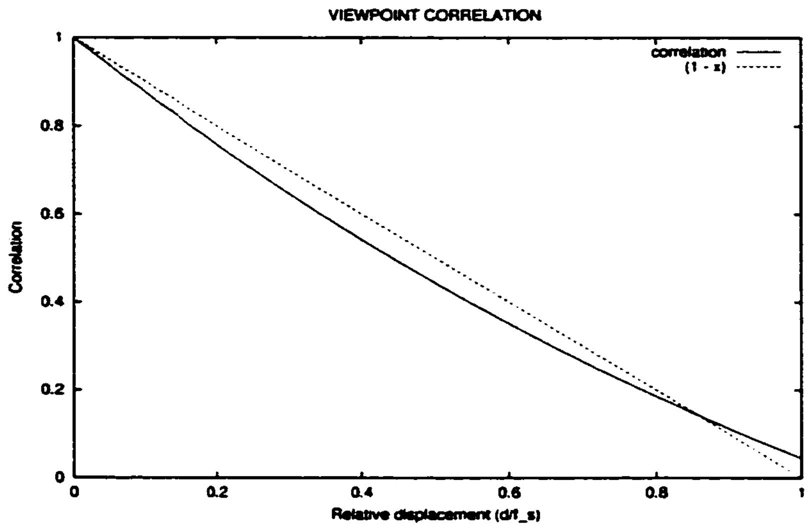


Figure 6.11: Correlation Approximation

$$\mu_{A_{rel}} \approx 1 - \frac{d}{f_s} \tag{6.3}$$

for $d < f_s$, which fits with intuition. In our simplified model of viewpoint correlation, average A_{rel} is a suitable approximation for viewpoint correlation over translation in the optimal scanning zone $\mathbf{S_o}$[6]. Therefore, we can approximate viewpoint correlation σ_v by

$$\begin{aligned} \sigma_v &\approx 1 - \frac{d}{f_s}, d \leq f_s \\ &= 0, d > f_s. \end{aligned} \tag{6.4}$$

[6]It is important to note that measures of distance and area in this simplified formulation are indirect measurements referenced back to the rough model.

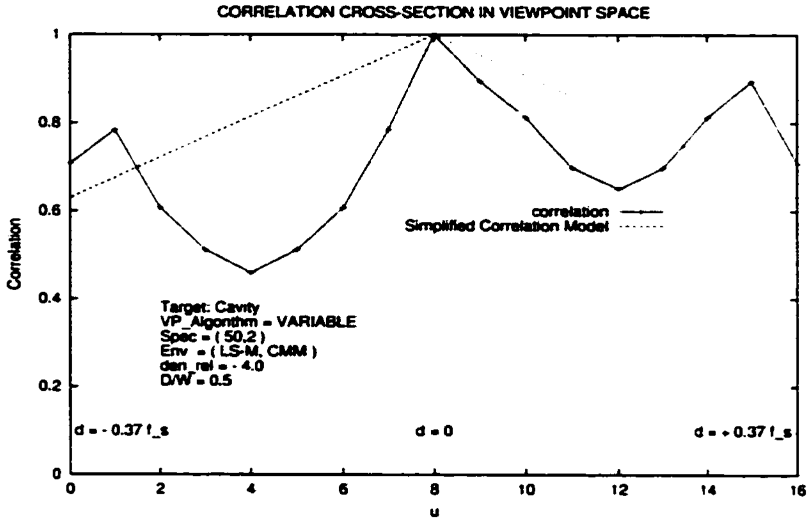


Figure 6.12: Correlation Cross-section - Cavity

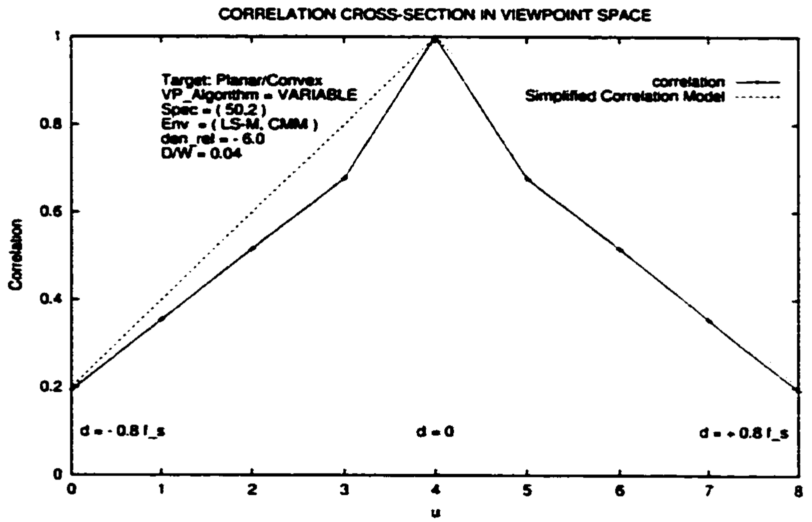


Figure 6.13: Correlation Cross-section - Planar/Convex

To illustrate the suitability of this simple model, we show profiles through the correlation data presented earlier at Figures 4.3 and 4.4. Figure 6.12 is a profile through the correlation surface[7] for the cavity target shown at Figure 4.4 for translation parameter $v = 8$. Similarly, Figure 6.13 is a profile through the correlation surface for the slightly convex target shown at Figure 4.3 for translation parameter $v = 4$. These figures show that the approximation is good for nearly planar targets but rather crude for more complex shapes. Nevertheless, we find that it is adequate to develop rules of thumb for viewpoint space sampling.

Assuming an even distribution of viewpoint positions on the optimal scanning zone $\mathbf{S_o}$, the positional density of viewpoints ρ_v can be expressed as

$$\rho_v = \frac{1}{d^2} \tag{6.5}$$

where d is the inter-viewpoint spacing on $\mathbf{S_o}$, referenced back to measurement between the corresponding vertices of the rough model[8]. Substituting the value of d in Equation 6.5 into the expression for viewpoint correlation σ_v at Equation 6.4 we get the following relationship between viewpoint correlation and viewpoint sampling density over $\mathbf{S_o}$:

$$\sigma_v \;\approx\; 1 - \frac{1}{f_s\sqrt{\rho_v}}$$

[7]Recall that the correlation surface was computed as the viewpoint correlation relative to the center vertex of the surface patch as a function of translation parameters u and v on the optimal scanning surface. A correlation profile is a slice through the correlation surface for a fixed translation parameter, in this case the parameter v.

[8]It is convenient to reference distances in viewpoint space back to the object surface as viewpoints are generated directly from the rough model.

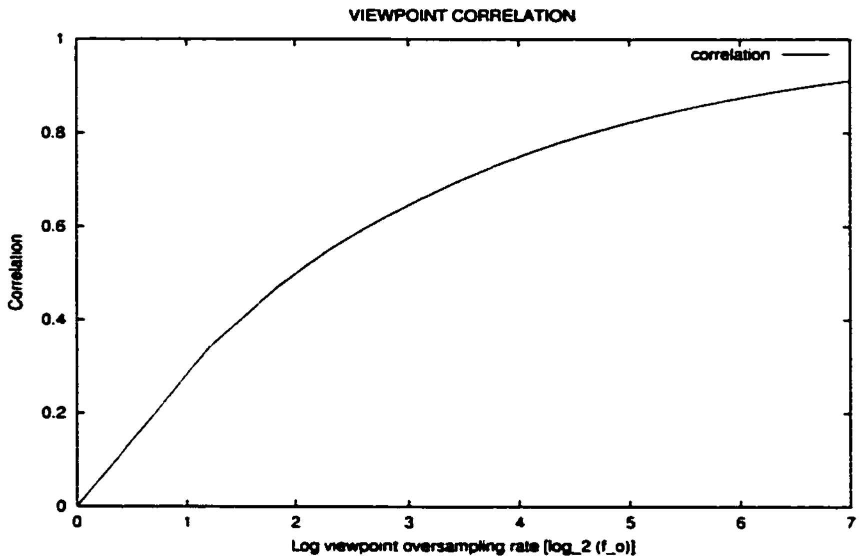


Figure 6.14: Correlation Versus Viewpoint Density

$$
\begin{aligned}
&= 1 - \frac{1}{\sqrt{F_s \rho_v}} \\
&= 1 - \frac{1}{\sqrt{f_o}}. \qquad (6.6)
\end{aligned}
$$

Then,

$$f_o \approx \frac{1}{(1-\sigma_v)^2}. \qquad (6.7)$$

The relationship is plotted at Figure 6.14. The units of ρ_v are viewpoints per unit area, where area is referenced back to the rough model. $f_o = F_s \rho_v$ is the over-sampling rate with respect to the position component of viewpoint space and is unit-less. For example, $f_o = 4$ provides a 50% overlap and an average viewpoint correlation of approximately 50%. Because of the many simplifying assumptions in the underlying models, Equation 6.6 and Figure

6.14 should not be taken too literally. Nevertheless, they provide useful insights into the relationship between the degree of oversampling in viewpoint space and viewpoint correlation. Our experiments show that, as a rule of thumb, a value of viewpoint correlation $\sigma_v = 0.75$ which corresponds to an over-sampling rate of $f_o = 16$ is a good initial choice, subject to fine tuning depending on the reconstruction problem. Also note that the relationship between viewpoint correlation and sampling density involves only translation over the optimal scanning zone $\mathbf{S_o}$. Oversampling in viewpoint orientation is a different matter.

6.4.2 Position-Orientation Tradeoffs

In Section 6.3 we introduced a viewpoint generation scheme based on an optimal viewpoint zone $\mathbf{S_o}$ tied to a rough model of the object. It was postulated that the set of viewpoints thus generated provides a near-optimal sampling of viewpoint space for the purposes of solving the view planning problem. With a one-to-one correspondence between surface points and candidate viewpoints, the complexity of generating the measurability matrix without segmentation became $\mathcal{O}(s^3)$.

It is useful to consider decoupling the discretization scheme in surface space S and viewpoint space V. In creating the rough model, the mesh has been progressively decimated to a level of geometric resolution just satisfactory for model-based view planning [SRR01c]. The decimation process concentrates vertex resources in regions of high curvature, an attribute desireable both for conserving shape fidelity and for viewpoint generation. The level of surface space discretization is driven primarily by the shape fidelity necessary for surface normal estimation and for visibility determination.

In discretizing viewpoint space, there are useful trade-offs between quantization levels in the position and orientation components of sensor pose.

This should allow a degree of decoupling between discretization levels in surface space and viewpoint space. While viewpoint correlation is strongly influenced by object shape, it is evident that there is a great deal of redundancy in viewpoint space. The degree of viewpoint correlation suggests we can afford a further sub-sampling of V by an additional decimation factor d_v - that is, reduce v to s/d_v.

A further issue to be considered is minimization of shadow effects which are sensitive to sensor twist orientation. An example is an elongated cavity feature. Experiments have shown that it is difficult to characterize such features sufficiently reliably and accurately to automatically set the twist orientation component. Pending development of sufficiently reliable surface segmentation and feature characterization algorithms, it will be necessary to separately discretize twist rotation. The number of quantization intervals d_t is to be experimentally determined. This has the effect of increasing the dimensionality of candidate viewpoint space from two to three. Combining this revised viewpoint space sampling strategy with segmentation, the overall measurability matrix complexity is $\mathcal{O}(\frac{s^3 d_t}{c^2 d_v})$.

As we wish to maintain a robust but efficient level of viewpoint correlation, it is useful to develop rules of thumb for the degree of sub-sampling possible. In Chapter 4 we noted that the minimum number of viewpoints required for a complete set covering is defined by the footprint ratio $r_f = F_p/F_s$, the ratio of the patch and sensor footprints. For a given viewpoint oversampling rate f_o, the number of viewpoints required is approximately $f_o r_f$. If we wish to consider a number of viewpoints v less than s, then we need

$$v \geq f_o r_f \tag{6.8}$$

viewpoints. For a further decimation of the rough model by the decimation ratio d_v, where

$$d_v = \frac{s}{v}, \tag{6.9}$$

then we need

$$\frac{s}{d_v} \geq f_o r_f. \tag{6.10}$$

For a segmented patch rather than a closed model surface, it is necessary to deal with edge effects, in which case the following is a better measure of the decimation ratio d_v than Equation 6.9

$$d_v = \frac{(\sqrt{s}-1)^2}{(\sqrt{v}-1)^2}. \tag{6.11}$$

Consequently, for a given viewpoint oversampling ratio f_o, we have an approximate limit on the viewpoint decimation ratio of

$$d_v \leq \frac{s}{f_o r_f}. \tag{6.12}$$

Similarly, for a given viewpoint decimation ratio d_v, the viewpoint oversampling rate is approximately

$$f_o \approx \frac{s}{d_v r_f}. \tag{6.13}$$

6.4.3 Decoupled Algorithm

The DECOUPLED algorithm (Figure 6.15) first decouples the level of viewpoint position discretization from the level of rough model surface discretization. It then further decouples the level of twist orientation discretization from the number of viewpoint positions.

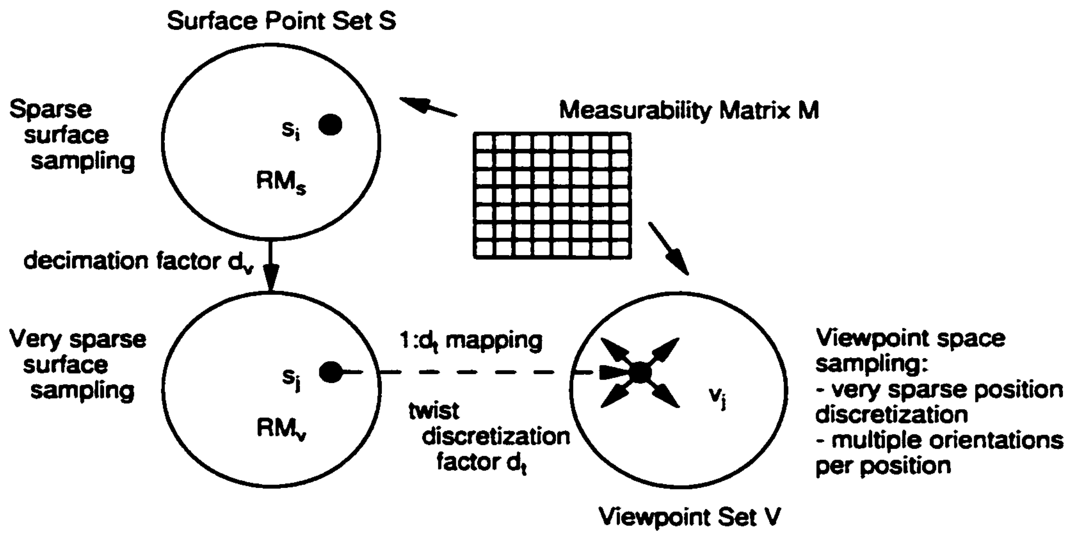


Figure 6.15: Viewpoint generation - Decoupled Algorithm

The essence of the DECOUPLED algorithm, then, is further decimation of the already coarse rough model for the purpose of efficient viewpoint generation. We now employ two rough model versions - one for representing surface geometry and the other for viewpoint generation. The original rough model RM_s is unchanged from our previous definition. It remains the basis of sampling object surface space and is tied to rows in the measurability matrix. RM_s is further decimated to create RM_v which becomes the basis for viewpoint generation and relates to columns in the measurability matrix.

The strategy maintains the original coarse rough model sampling in RM_s at a level of shape fidelity sufficient for visibility analysis and surface normal estimation. Very sparse sampling in the position component of viewpoint space in RM_v reduces viewpoint correlation to a more efficient level while increased sampling of the twist component of orientation is more robust with respect to object shape complexity and to shadow effects[9].

To evaluate the DECOUPLED view planning strategy, a series of experiments were conducted on the slightly convex target used previously in Chapter 4. View planning performance data using the VARIABLE algorithm for that target was shown in Chapter 4, Section 4.4.2 at Figures (4.9-4.14). Experimental data on the same target using the DECOUPLED algorithm is reported at Table 6.1. Here, d_t/d_v is a measure of the relative complexity of the DECOUPLED versus the VARIABLE algorithms; n_{Best} is the most efficient view plan length for the candidate viewpoint set as found by exhaustive search; and n is the view plan length with the greedy search algorithm. Recall from Chapter 4 that the optimum view plan length for this target is $n_{Opt} = 4$, as determined by inspection. Additionally, greedy search with the VARIABLE algorithm gave a view plan of length $n = 7$ while exhaustive search on the same candidate viewpoint set gave a best view plan length of $n_{Best} = 6$.

It is interesting to apply expressions derived from the simplified correlation model to the two data sets presented at Table 6.1. With a footprint ratio $r_f = 2.55$, sensor footprint $F_s = 2f_d R_o T \Phi_2 L_y = 6145.03$ and rough model size $s = 289$, we get the values for over-sampling rate f_o, average viewpoint correlation σ_v and viewpoint positional density ρ_v as shown in Table 6.3.

The DECOUPLED experimental data is summarized at Table 6.2 and Figure 6.16, where we have averaged the raw performance data. We note that the DECOUPLED algorithm achieves optimum view planning efficiency for this data set with a value of relative complexity d_t/d_v in the range $[0-1]$.

[9]Note that some range scanners seek to achieve these goals by using multiple receive heads (usually two) for a single laser head. Such a configuration amounts to multiple sampling of viewpoint space in twist angle for a single physical sensor pose.

d_v	d_t	$d_t/\ d_v$	$\log_2(d_t/\ d_v)$	δt	n_{Best}	n	$n_{Opt}/\ n_{Best}$	$n_{Opt}/\ n$
4	1	1/4	-2	360°	6	8	0.67	0.5
4	2	1/2	-1	180°	6	8	0.67	0.5
4	4	1	0	90°	5	5	0.8	0.8
4	8	2	1	45°	4	5	1.0	0.8
4	16	4	2	22.5°	4	6	1.0	0.67
4	32	8	3	11.25°	4	6	1.0	0.67
16	1	1/16	-4	360°	6	9	0.67	0.44
16	2	1/8	-3	180°	6	9	0.67	0.44
16	4	1/4	-2	90°	6	8	0.67	0.5
16	8	1/2	-1	45°	4	5	1.0	0.8
16	16	1	0	22.5°	4	5	1.0	0.8
16	32	2	1	11.25°	4	5	1.0	0.8

Table 6.1: View Plan Efficiency - Decoupled Algorithm - Convex Patch

$\log_2(d_t/\ d_v)$	$\overline{n_{Opt}/n_{Best}}$	$\overline{n_{Opt}/n}$
-4	0.67	0.44
-3	0.67	0.44
-2	0.67	0.5
-1	0.8	0.62
0	0.88	0.8
1	1.0	0.8
2	1.0	0.67
3	1.0	0.67

Table 6.2: Relative View Plan Efficiency - Decoupled Algorithm - Convex Patch

Data set	d_v	f_o	$\log_2 f_o$	σ_v	ρ_v
1	4	28.3	4.82	0.81	0.0046
2	16	7.08	2.82	0.62	0.0012

Table 6.3: Viewpoint Sampling Parameters - Decoupled Algorithm - Convex Patch

At these decimation levels in viewpoint position and twist angle, the complexity of the DECOUPLED and VARIABLE algorithms is nearly identical, yet performance has improved. The data suggests that the DECOUPLED algorithm will give good results with a viewpoint decimation factor d_v of about $[4 - 8]$ and twist decimation factor d_t of about $8 - 16$, that is to say twist increments of about $45° - 22.5°$.

6.5 Conclusions

In this chapter, we have presented the concept of model-based viewpoint generation and the optimal scanning zone $\mathbf{S_o}$. Based on that concept, we presented two near-optimal viewpoint generation algorithms - the "optimal scanning zone" and the "decoupled" algorithms. The DECOUPLED algorithm was shown to have comparable computational complexity but improved performance and robustness to object shape complexity. By exploiting knowledge of the sensor and the approximate object geometry as represented by the rough model, we are able to generate viewpoints which are individually optimal, and collectively near-optimal. We examined the issue of over-sampling in viewpoint space through the means of a simplified viewpoint correlation model which permitted formulation of some rough rules of thumb for viewpoint sampling.

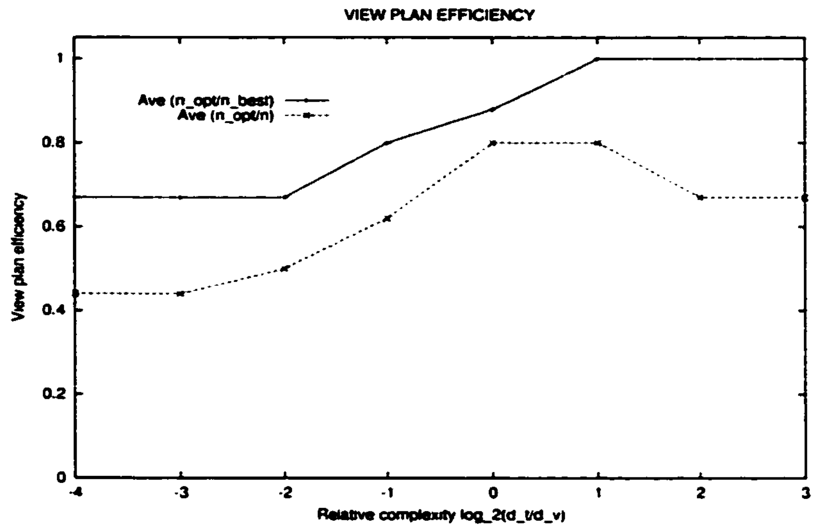


Figure 6.16: View Plan Efficiency - Decoupled Algorithm

With respect to efficiency, we note that the algorithm is inherently parallelizable. For the inspection application, the measurability matrix can be computed once off-line and inspection plans generated on-line. For object reconstruction, a combination of sparse sampling in object and viewpoint space, segmentation, parallelization of measurability matrix computations, and a greedy search set covering algorithm combine to offer good view planning performance within acceptable time limits for industrial applications.

Accuracy and robustness have been addressed by several measures. Firstly, the method includes a comprehensive model of sensor performance, including the most important error mechanisms. Sampling of object surface space and viewpoint space is sparse yet robust to errors. Greedy search provides a simple and robust set covering algorithm.

Aspects of view planning which are missing from the current work are the ability to deal with specular surfaces or other artifacts such as corner reflections, geometric and reflectance edge effects. These aspects are in the category of future work and require a more sophisticated model of sensor dynamic range and revised viewpoint generation and evaluation strategies.

Chapter 7

View Planning with Pose Uncertainty

This chapter examines the impact of pose error on view planning performance. View planning counter-measures are presented.

7.1 Introduction

We examine the effects on a single planned view of pose error resulting from positioning system inaccuracies. While the focus of the current work is on model building, the analysis is equally applicable to the inspection application. The performance of the positioning system impacts the view planning process, regardless of whether a traditional or performance-oriented view planning algorithm is utilized. For the purposes of this analysis, pose errors are broken down into position error, sensor boresight pointing error (axis error) and rotation error about the sensor boresight (twist error). For each error type, we examine its impact on measurement precision, sampling density, visibility and frustum occupancy. Pose error effects are analyzed in detail for one specific, common sensor configuration, the line-scan range camera. Results are generalizable to other range camera geometries. Input

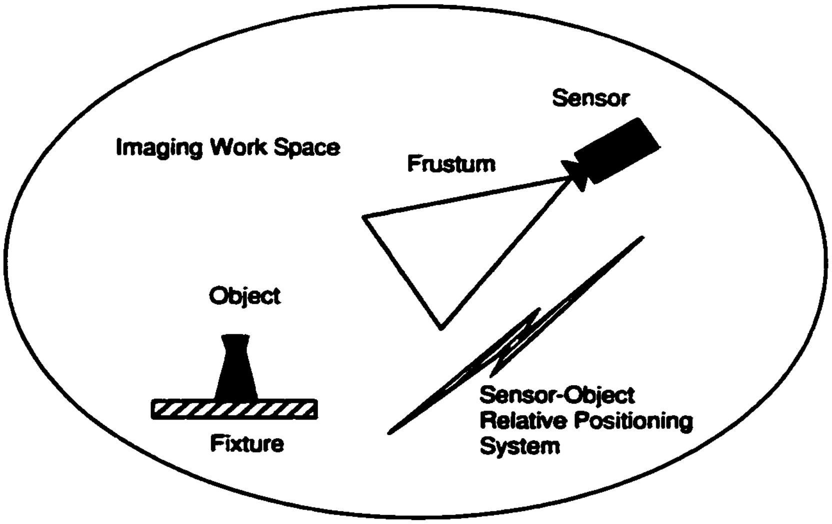


Figure 7.1: The Geometric Imaging Environment

and output parameters in the problem description are normalized and the results quantified, where possible.

An outline of the analysis is as follows. We begin in section 7.2 by defining performance criteria, briefly examine the overall effect of pose error on view planning, provide an overview of positioning systems, define a simple pose error model and define sensor geometry. Section 7.3 outlines the analysis methodology, summarizes results in tabular and graphical form, analyses the issues and discusses means to mitigate the effects of pose error on view planning. Mathematical details are contained in Appendix C. A summary and conclusions are presented in Section 7.4.

It is anticipated that the analysis may benefit imaging system design (specifying compatible sensor and positioning system performance) and view planning algorithm design (compensating for positioning system inaccuracy).

7.2 The View Planning Context

7.2.1 Performance Criteria

Performance-oriented view planning incorporates two inputs not found in traditional view planning approaches - a model specification and an imaging environment specification. The model specification defines reconstruction goals, such as measurement precision and sampling density, which may be fixed or variable over the object surface. The imaging environment specification defines key sensor and positioning system model parameters.

As noted in Section 3.3, for a surface point to be declared measurable, all model specification requirements must be met at that surface point for that viewpoint. Specifically, the following tests must be passed:

- *frustum occupancy* - the surface point must fall within the sensor frustum for that viewpoint,
- *visibility* - the surface point must be locally visible from the optical source and receiver positions defined by that viewpoint, and
- *sampling precision and density* - the estimated sampling precision and sampling density at the surface point for that viewpoint must meet the specified requirements.

7.2.2 Pose Uncertainty Effects

Unfortunately, positioning system errors negatively impact all of these requirements. View planning is a computationally-intensive task with the objective of arriving at a small set of optimal or near-optimal viewpoints, the next-best-view (NBV) list. When the NBV list is sent to a positioning

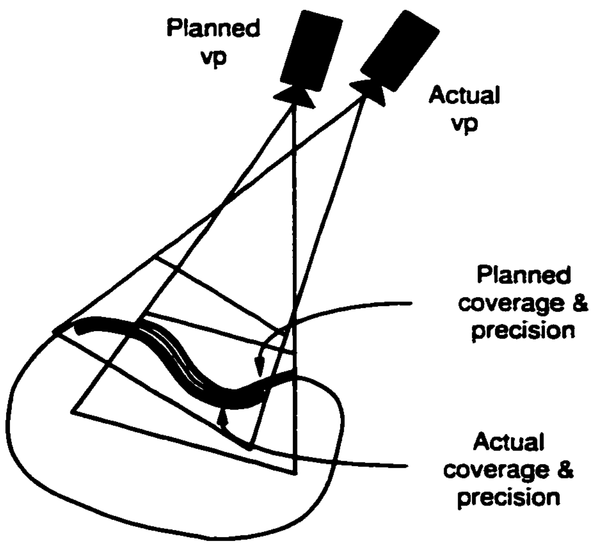


Figure 7.2: View Planning with Pose Uncertainty

system whose position and orientation accuracy is inferior to that of the sensor, the coverage of individual viewpoints and of the NBV list as a whole is compromised. Individual viewpoint positions and orientations are corrupted. Orientation error is particularly troublesome as effects are amplified by range. As illustrated in Figure 7.2, image coverage (frustum occupancy), measurement precision and sampling density will all be effected. Visibility can also be effected by the altered viewing geometry.

We can recover a refined pose estimate post-facto by employing suitable registration techniques and subsequently re-estimate measurement quality within the acquired image. However, we are still left with data acquisition differing from that which had been planned. As pose error deteriorates, the computationally intensive view planning phase is progressively compromised - ultimately to be rendered futile. Consequently, there is a need to make the view planning process robust with respect to pose uncertainty resulting from positioning system errors.

Before we can devise compensatory measures, it is necessary to quantify and qualify the various effects. We begin our analysis by presenting a generalized model for positioning system error.

7.2.3 Positioning Systems

A variety of positioning systems are in common usage, covering a wide range of accuracy. These include co-ordinate measuring machines (CMMs), translation stages, turntables and other rotary joints as well as robot arms and similar devices[1]. At the top end, CMMs offer accuracy superior to the best range camera. At the other extreme, robot arms and similar devices provide good repeatability but comparatively poor accuracy relative to high quality range cameras.

In addition to positioning and orienting the sensor in space, positioning systems are often used to provide one dimension of the sensor scan. For example, a laser scan in the camera x-z plane may be swept along the camera y-axis (line-scan mode) or rotated about an axis parallel to the camera x-axis (cylindrical-scan mode). In such cases, positioning system performance directly effects not only what is scanned (by setting the camera pose) but also the quality of measurement within the range image (by the degree of mechanical jitter). For the present work, we assume that optical and mechanical system components have compatible performance at the image pixel level and restrict our examination to the effects of pose error on overall image coverage and measurement quality.

[1] A good tutorial on positioning system error mechanisms can be found at [New00].

7.2.4 Positioning System Error Model

In general, it is difficult to characterize the accuracy of positioning systems with multiple degrees of freedom [SW92], [STS92], [CMKD96]. Accuracy can also be highly variable over the movement envelope for a given machine. For the purposes of analysis, therefore, we adopt the following simplified but general purpose pose error model. First, we assume that calibration has identified, measured and removed from further consideration all systematic error components, leaving only the residual stochastic errors.

Errors in sensor position, boresight axis and rotation about the boresight (twist) are considered to be independent random processes. Position error is modeled as a zero-mean Gaussian process with standard deviation σ_p uniformly distributed in off-set direction. Axis error is modeled by a unit vector uniformly distributed on the surface of a cone centered on the camera boresight where the cone half-angle is a zero-mean Gaussian process with standard deviation σ_a. Twist error is modeled as a zero-mean Gaussian process with standard deviation σ_t. In order to gain an appreciation of the mechanisms involved and their relative importance, we will separately examine each pose error effect in isolation.

While the model just described is a suitable general purpose framework for analyzing the effects of positioning system error, in practice it will be necessary to develop and apply a specific error model tailored to the type, configuration and movement envelope of each unique positioning system in actual usage.

7.2.5 Range Camera Geometry

To illustrate pose error effects, we examine the case of a line-scan range camera, a common configuration whose imaging geometry was previously shown

	Position	Axis	Twist
Measurement Precision	-	-	-
Sampling Density	-	-	-
Visibility	-	-	-
Frustum Occupancy	-	-	-

Table 7.1: Pose Error Effects

at Figure 3.7. Variables were defined in Section 3.4.1. A line-scan range camera takes range measurements by sweeping the laser in the camera x-z plane. The entire camera is then mechanically swept along the camera y-axis for a specified distance. Pose error effects with other imaging configurations depend on the measurement geometry but are similar to those presented here for the line-scan range camera case.

7.3 Pose Error with a Single View

This section analyzes the effects of pose error on a single view (Figure 7.3). We separately consider errors in position and orientation (camera axis and twist about the camera axis). For each of these, we examine the impact on measurement precision, sampling density, visibility and frustum occupancy to complete a table of pose error effects as illustrated at Table 7.1. Results are summarized at the end of the section.

7.3.1 Performance Models

We begin by defining models for relative measurement precision, sampling density and frustum occupancy. Visibility will be addressed in the next section.

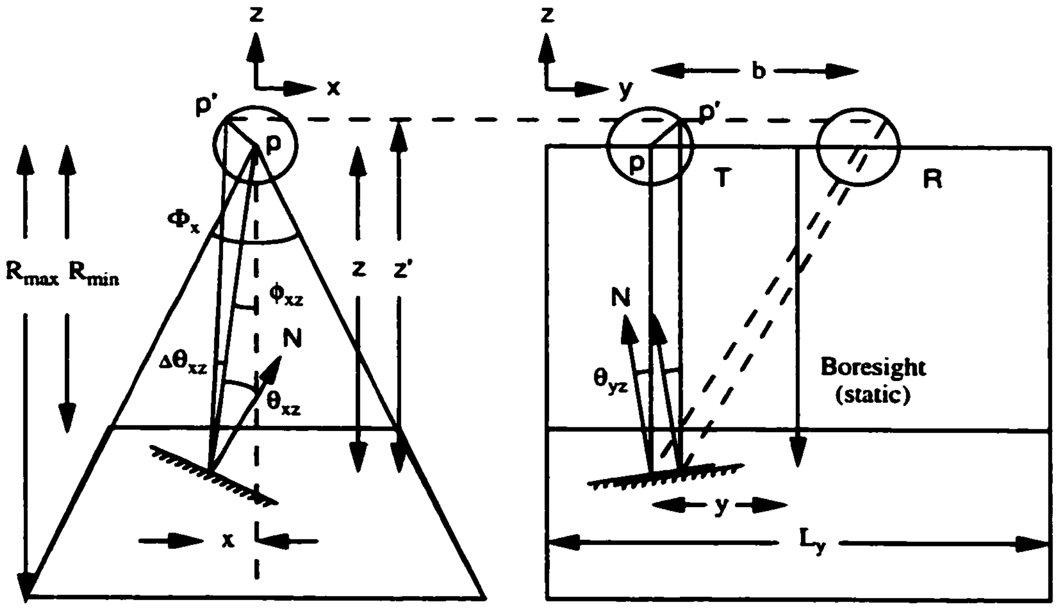


Figure 7.3: Line-scan Range Camera Geometry

Measurement Precision

In Section 3.4.2 and Equation 3.4, we derived the following model of sensor measurement precision $\hat{\sigma}_z$ -

$$\hat{\sigma}_z = \frac{C_z z^2}{\cos\theta_{yz}\ U_i(\theta_{yz} - t_{yz})\ U_i(\theta_{xz} - t_{xz})}. \tag{7.1}$$

In the foregoing, the x,y,z axes define the sensor coordinate frame, θ_{xz} and θ_{yz} are the inclination angles of the surface element with respect to the laser in the x-z and y-z planes, t_{xz} and t_{yz} are inclination angle thresholds, $U_i(\theta)$ is the inverse unit step function and C_z is the z-axis geometric noise coefficient derived from sensor calibration.

An inherent feature of optimal candidate viewpoint generation is that inclination angles are zero for the targeted surface point and are small for its immediate surrounding region. Additionally, the level of pose error under consideration results in only small perturbations to observed inclination angles. Therefore, we are justified in removing inclination angle threshold effects from Equation 7.1 to get the following simplified expression for the standard deviation of the estimated geometric noise component along the sensor z-axis:

$$\hat{\sigma}_z = \frac{C_z z^2}{\cos\theta_{yz}}. \tag{7.2}$$

To consider the relative impact of pose error on measurement precision, it is convenient to define relative precision P_{rel} as the ratio of estimated precision in the case of pose error to estimated precision in the error-free case. Then,

$$P_{rel} = \frac{\hat{\sigma}_{z'}}{\hat{\sigma}_z} \tag{7.3}$$

which reduces to

$$P_{rel} = \frac{z^2}{z'^2} \frac{\cos\theta'_{yz}}{\cos\theta_{yz}} \tag{7.4}$$

where z' and θ'_{yz} are the range and y-z incidence angle as perturbed by pose error.

Sampling Density

In Section 3.4.3 and Equation 3.8, we derived the following model of sampling density $\hat{\rho}_z$:

$$\hat{\rho}_z = \frac{(N_x - 1)^2 (N_y - 1)^2 C^2\theta_{xz} C^2\theta_{yz}}{R_{xz}{}^2 \Phi_x{}^2 (N_y - 1)^2 C^2\theta_{yz} + L_y{}^2 (N_x - 1)^2 C^2\theta_{xz}}. \tag{7.5}$$

Here, image size is N_x-by-N_y samples, Φ_x is the sensor field of view in the scanning plane, R_{xz} is the slant range, L_y is the linear scan length in the y-z plane, ϕ_{xz} is the instantaneous laser scan angle in the x-z plane, and θ_{xz} and θ_{yz} are the inclination angles in the x-z and y-z planes. As noted previously, we adopt the abbreviated notation $C\theta = \cos\theta$ for simplicity.

In most cases, Equation 7.5 can be simplified for symmetric range images by setting $N = N_x = N_y$, giving

$$\hat{\rho}_z = \frac{(N - 1)^2 C^2\theta_{xz} C^2\theta_{yz}}{R_{xz}{}^2 \Phi_x{}^2 C^2\theta_{yz} + L_y{}^2 C^2\theta_{xz}}. \tag{7.6}$$

Further, a well-designed viewpoint will set $\delta x = \delta y$ for the optimal sampling geometry. This sets stand-off range $z = f_d R_o$ (where R_o is the optimum sensor scanning range and f_d is a stand-off distance factor $f_d = 1 + \delta, \delta \ll 1$), scanning angle $\phi_{xz} = 0$ and inclination angles $\theta_{xz} = \theta_{yz} = 0$. This is achieved by adjusting the linear scan length L_y such that $L_y = f_d R_o \Phi_x$. Consequently, for an optimized viewpoint, Equation 7.6 becomes

$$\hat{\rho}_z = \frac{(N - 1)^2 C^2\theta_{xz} C^2\theta_{yz}}{\Phi_x{}^2 [R_{xz}{}^2 C^2\theta_{yz} + f_d{}^2 R_o{}^2 C^2\theta_{xz}]}. \tag{7.7}$$

As before, to consider the relative impact of the position component of pose error on sampling density, it is convenient to define relative sampling density D_{rel} as

$$D_{rel} = \frac{\hat{\rho}_{z'}}{\hat{\rho}_z}. \tag{7.8}$$

Then,

$$D_{rel} = \frac{C^2\theta'_{xz}C^2\theta'_{yz}}{C^2\theta_{xz}C^2\theta_{yz}} \frac{R_{xz}{}^2C^2\theta_{yz} + f_d{}^2R_o{}^2C^2\theta_{xz}}{R'_{xz}{}^2C^2\theta'_{yz} + f_d{}^2R_o{}^2C^2\theta'_{xz}}. \tag{7.9}$$

Frustum Occupancy

Frustum Erosion As illustrated in Figure 7.4, pose uncertainty reduces the portion of the frustum that will confidently image a given spatial volume. Conversely, if we consider the frustum as a spatial probability density function (pdf) representing the probability of the sensor sampling a given spatial region, then positioning system uncertainty has the effect of blurring the otherwise crisp pdf shape and narrowing the high confidence sampling zone. The actual frustum remains unchanged, of course. Pose position error impacts both the effective depth of field (DOF) and effective field of view (FOV). Pointing uncertainty mainly impacts the effective field of view. Twist error only impacts the effective field of view.

Frustum erosion effects are non-linear and involve interaction between the frustum shape (related to sensor type) and positioning system characteristics. The illustration at Figure 7.4 depicts a line-scan range camera in a robotic "eye-in-hand" configuration. Each imaging environment is unique and must be separately modeled. However, all share the common characteristic of erosion of the frustum confidence zone.

Field of View Erosion We define a viewpoint's targeted footprint TF as the frustum cross-sectional area at the stand-off range for that viewpoint. The sensor footprint remains unchanged with viewpoint perturbation but pose error will cause it to cover some unplanned regions while losing some planned coverage. The unplanned coverage gain is of no direct benefit in

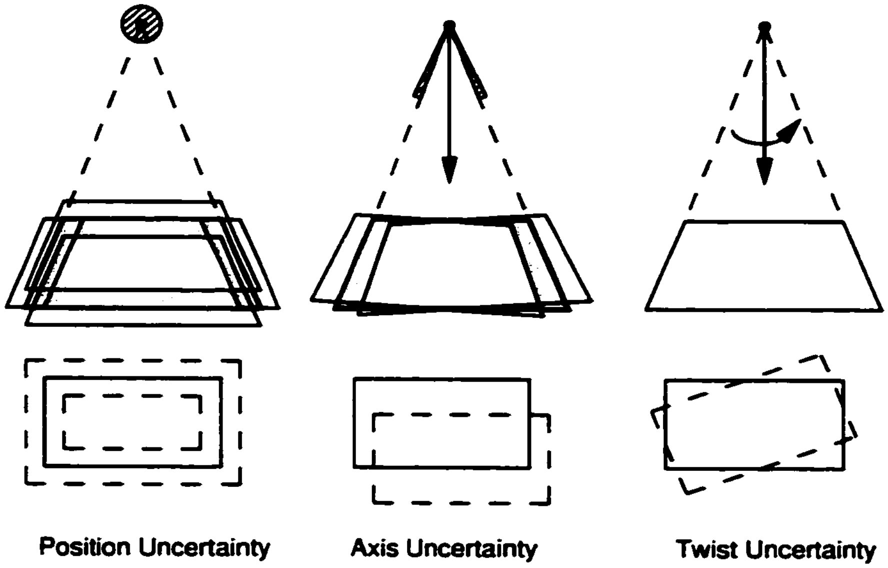


Figure 7.4: Frustum Erosion with Pose Uncertainty

planning a single view while the unplanned coverage loss erodes view planning effectiveness. Thus, for view planning purposes, pose error always reduces and never increases effective coverage.

For a line-scan sensor, the targeted footprint is

$$TF = \underbrace{(2f_d R_o T \Phi_2)}_{width} \underbrace{(L_y)}_{length} \tag{7.10}$$

where $\Phi_2 = \Phi_x/2$, that is - half the field of view in the x-z plane.

To consider the relative impact of the position component of pose error on effective frustum field of view, it is convenient to define the relative targeted footprint TF_{rel} as

$$TF_{rel} = \frac{TF'}{TF} \tag{7.11}$$

where TF' is the portion of the sensor footprint disturbed by pose error which overlaps the targeted footprint.

Depth of Field Erosion In considering depth of field, for all practical purposes, only erosion of the effective near-field range limit is of concern. For optimum measurement performance, the stand-off range is set at $z = f_d R_o = f_d R_{min}$. Pose error will result in erosion whenever the z-component of frustum change $\delta_z > (f_d - 1)R_o$. Therefore, the statistic of interest relative to depth of field erosion is the probability $P[\delta_z > (f_d - 1)R_o]$.

7.3.2 Analysis Methodology

Briefly, the methodology used in the analysis is as follows. The effect of each type of pose error on each performance criteria (precision, sampling density, visibility and frustum occupancy) is separately studied. Performance in the presence of pose error is normalized to that for the error-free case. In each case, the analysis begins by applying error effects to the models for estimation of measurement precision, sampling density and sensor coverage, as defined in the previous section. Next, the statistics of key functions of the underlying random variables are computed. These are used to compute the expected value μ and variance σ^2 for the relative performance measure. Mathematical details of the pose error analysis are contained in Appendix C.

7.4 Summary and Conclusions

7.4.1 Overview

Based on the analysis at Appendix C, Table 7.2 provides a qualitative overview of pose error effects on a single view. On a relative basis, effects are broadly

	Position	Axis	Twist
Measurement Precision	L-M	L	nil
Sampling Density	L-M	L	nil
Visibility	L	L	M
Frustum Occupancy	M-H	H	L

Table 7.2: Qualitative Pose Error Effects

categorized in the range: high (H), moderate (M), low (L) or nil. Analytical results in pose error variables up to second order effects are presented in Tables 7.3, 7.4 and 7.5 at the end of this chapter.

Also located at the end of this chapter are a series of figures (Figures 7.5 - 7.11) which plot error effects with non-zero variance, allowing a quantitative comparison. Briefly, the figures can be interpreted as follows. More detailed interpretation of the figures can be found in Appendix C.

- Figures 7.5 and 7.6 - The ordinate is relative precision P_{rel} and relative sampling density D_{rel}, respectively. The abscissa is normalized position error σ_p/R_o. One- and two-sigma curves for relative precision (or sampling density) are displayed bracketing the nominal value which is close to 1.0. Precision improves for values $P_{rel} < 1$, while sampling density improves for values $D_{rel} > 1$. Any vertical slice through these curves defines the shape of the probability density function for relative precision (or sampling density) for a given normalized position error.

- Figures 7.7 and 7.8 - The ordinate is the logarithm base ten of the probability of depth of field erosion. The abscissa is normalized position error σ_p/R_o and axis orientation error σ_a, respectively. Probability curves are plotted for various levels of sensor stand-off distance factor f_d.

- Figures 7.9, 7.10 and 7.11 - The ordinate is relative targeted footprint TF_{rel} while the abscissa is normalized position error σ_p/R_o, axis orientation error σ_a and twist orientation error σ_t, respectively. They can be interpreted in the same manner as the first pair of figures. Relative targeted footprint TF_{rel} is always ≤ 1.0.

The effects of pose position and pose orientation error are shown. The latter is broken down into sensor boresight axis error and rotation error about the sensor boresight (twist). In all cases, performance is relative to the error-free case.

Finally, note that the foregoing results are based on an optimized scanning configuration of a line-scan range camera, the conditions for which are summarized in Table 7.6. Results for non-optimized scanning configurations can be found in the appropriate sub-sections of Appendix C.

7.4.2 Pose Error Model Suitability

This analysis has employed a simple, general purpose error model based on Gaussian statistics. The statistical behavior of real positioning systems is more complex and highly variable between system type and configuration. Pose error is also very difficult to characterize and quantify in practice. Furthermore, each type of pose error effect has been studied in isolation. Multiple types of pose error in combination will interact in a non-linear manner. Nevertheless, while its limitations are acknowledged, the pose error model employed here should form a suitable basis for understanding pose error effects and in selecting countermeasures to deal with them.

7.4.3 Measurement Performance

The effects of pose error on measurement performance (precision and sampling density) for a single viewpoint are summarized at Table 7.3. Average measurement performance is only slightly effected by second order effects which will be small in most practical cases.

In the case of position error, average relative measurement precision is slightly degraded due to the quadratic variation of sensor precision with range. On the other hand, average relative sampling density is slightly improved as the improvement at shorter ranges slightly outweighs degradation at longer ranges. However, the important statistic is measurement variance which is low to moderate. Unless taken into account, pose position error will degrade view planning performance.

For axis orientation error, average relative measurement precision is marginally improved as range to the target is reduced. Average relative sampling density is slightly degraded by off-boresight and inclination angle effects. Measurement variance is zero as a result of symmetry. Thus, the net effect of axis orientation error on measurement performance is low.

For twist orientation error, both average relative measurement precision and average relative sampling density are unchanged and the measurement variance is zero due to symmetry. Considered in isolation, twist orientation error has negligible effect on measurement performance. However, visibility problems may arise with twist error.

7.4.4 Visibility

Visibility or occlusion effects depend on several variables, only a few of which are restricted to the sensor or positioning system. The most important sensor parameter is the length of the optical baseline. Fixed for any given sensor, this parameter directly influences measurement performance, shadow effects and the ability to observe the interior of cavities. It is important to select a range camera with an optical baseline appropriate to the modeling task.

Two other major factors impacting visibility are object shape and relative sensor-object pose. The former cannot be altered and varies from task to task. Selecting appropriate sensor poses is the essence of view planning.

Once a pose has been generated, selected and included in the next-best-view list, pose error induced by the positioning system will influence actual visibility. In the case of pose position error, the visibility impact will be low to nil. The exception could be the case of surface regions being viewed at incidence angles near threshold or close to occlusion. This effect can be reduced by using conservative visibility tests in measurability calculations.

Axis orientation error has the effect of causing a wobble in the position of the optical transmitter and receiver. The degree of wobble is determined by the orientation error, the length of the optical baseline and the position of the scan along the y-axis. Effects are similar to those generated by pose position error. Visibility impacts will generally be low for small axis orientation error.

On the other hand, twist orientation error has the potential to significantly degrade the reliability of visibility predictions as it results in leveraged movement of the laser source and optical detector in a plane tangential to the targeted object region. Very small twist errors will have a minor impact in most imaging situations. However, moderate twist error can introduce significant occlusion problems with convoluted object shapes. This is also the reason why optimizing the twist component of pose is so important during generation of candidate viewpoints.

7.4.5 Frustum Occupancy

The third component of pose error is its effect on frustum occupancy - that is, sensor coverage. This is particularly problematic as the whole point of view planning is to generate an optimal set of NBVs with specific coverage and measurability. If actual viewpoint coverage is significantly altered, the entire view planning process is undermined and can be rendered futile.

Pose error erodes the reliable coverage zone i.e. the field of view and depth of view. The effects of pose error on field of view erosion for a single viewpoint are summarized at Table 7.4. Recall that Φ_x is the sensor field of view in the x-z plane and $\Phi_2 = \Phi_x/2$. Field of view erosion effects are shown at Figures 7.9, 7.10 and 7.11 for pose position, axis and twist error, respectively. From these curves, we observe that all three error components have a non-negligible impact on sensor coverage. We can characterize the effects as low for twist error, high for axis error and moderate-to-high for position error. As expected, axis error is particularly troublesome as the abbe effect amplifies coverage perturbations with stand-off range.

Some depth-of-field erosion occurs with pose position error. This is easily dealt with by a careful choice of the stand-off distance factor f_d. Depth-of-field erosion is negligible for pose orientation error.

7.4.6 Multiple Views

While the focus of the current work is on the effect of pose error on a single view, it is appropriate to briefly address the ramifications of pose error on view plans consisting of multiple views.

In most practical imaging situations, multiple view plans involve a great deal of redundant coverage for a variety of reasons:

- *Shape-driven image overlap* - The more complex the geometry, the greater the number of views required to capture the shape. This naturally leads to considerable image overlap around regions of high shape complexity.

- *Image-based registration overlap* - If the pose uncertainty is sufficiently large to necessitate image-based registration, then the registration constraint requires image overlap, which builds redundancy into the view plan.

- *Non-optimal set covering algorithms* - The set covering problem is known to be NP-complete. Thus, any practical set covering algorithm is sub-optimal, resulting in a view plan longer than the theoretical lower bound. Longer view plans equate to redundancy. Ironically, efficient view plans are more vulnerable to pose error effects.

- *Coverage gain from pose error* - This analysis has focussed on coverage loss due to pose error with a single view. However, where there is loss, there is gain. With multiple views, some coverage gain may partially off-set coverage losses.

The advantage of the redundancy phenomena is that it can mask most effects of low levels of pose error in some cases. The phenomena is not amenable to easy analytical treatment due to the large role played by object shape and the fact that view plans are typically small, such that composite coverage statistics are not well behaved.

The Tsimshian stone mask (Figure 6.5) is again an interesting object to illustrate the effects of pose error on a view plan consisting of multiple views. A view plan of five views was computed by the 3M technique for the back of the mask. The region poses a challenging scanning task due to the steep cavity walls and presence of several smaller cavities within the segment.

Performance of the view plan in the presence of multiple view pose error was evaluated by the experimental process described in Figure 4.7.

Figures 7.12, 7.13, 7.14 and 7.15 show pose error effects for positioning, axis and twist orientation errors plus their cumulative effect in combination. Average performance and error bars for +/- one standard deviation are shown. Partial-to-complete masking of pose error effects can be noted at low error levels, followed by a rapid decrease in average measurability and rapid increase in measurability variance with pose error deterioration. The data illustrates that, at low pose error levels, range scanning benefits from view plan redundancy while moderate to high levels of pose error result in unpredictable scanning performance. As the penalty for coverage failure is typically high for object reconstruction tasks[2], the unpredictable impact of pose error on multiple view sets may be unacceptable. Effects with simpler object shape will be more severe due to the lower level of viewpoint correlation.

7.4.7 Pose Error Countermeasures

A variety of countermeasures are available to deal with positioning system error. The quantitative analysis of this thesis may assist in setting key parameters.

- *Avoid the problem* - Pose error effects are sufficiently deleterious that they are best avoided by using the highest quality positioning system affordable. Mounting the range camera on a CMM or other high accuracy positioning system avoids many problems.

[2]The imaging team may have left the site or the object may no longer be readily available by the time coverage gaps or measurability deficiencies are discovered.

- *Compatible system design* - The specification and design of an object reconstruction system should treat the camera and positioning system as equally important components of an integrated system. Trade-offs in one impact the other.

- *System modeling and calibration* - Both the range camera and positioning system must be regularly calibrated. Good quality system models are essential and parameters should be updated after each calibration.

- *Optimize stand-off range selection* - The stand-off range can be selected to avoid depth of field erosion while optimizing measurement performance.

- *Robust measurability estimation* - Measurability estimation can be made robust to pose uncertainty by detuning compliance tests for frustum occupancy[3], specification compliance and visibility[4] during measurability matrix computations. By frustum detuning, we mean shrinking the effective frustum size. The amount of frustum detuning appropriate for a given level of pose error can be determined from Tables 7.4 and 7.5 and their associated figures. Specification compliance detuning means adding a compliance buffer by imposing somewhat more stringent tests on precision and sampling density than called for in the model specification. The size of compliance buffer required can be determined from Table 7.3 and the associated figures for a given level of pose error. This will, for example, make allowance for surface normal estimation error. Near-miss visibility testing applies to rays which are

[3]Tarabanis et al [TTA95] use a synthesis approach for generalized viewpoints which seeks to centralize viewpoints in the admissible domain.

[4]Tarbox [TG95] uses morphological erosion of viewpoints on the periphery of viewpoint sets to reduce vulnerability to pose error.

tangential or nearly tangential to the model surface. In these cases, we can increase the closest point of approach margin for declaring that an intersection has occurred. All of these measures provide increased margins for pose error. However, there are clearly practical limits beyond which view planning is rendered futile.

- *Add image-based registration constraint [SRR01a]* - This has similar effects to frustum detuning. The two constraints can be combined.

- *Feed back control loop* - Most positioning systems are open loop - that is, they are commanded to move to a specific pose and there is no control mechanism to confirm the fidelity with which the command was executed. It may be appropriate to consider a closed loop control system to reduce residual pose errors, perhaps through the use of machine vision or photogrammetric adjuncts to the positioning system.

Many of the foregoing countermeasures result in lengthened view plans. This has associated costs in scanning, registration and integration time and resources. Consequently, there are further trade-offs to be made between all phases of the object reconstruction process.

7.4.8 Conclusions

Performance-oriented view planning commences with a model specification requiring views (range images) to pass specific criteria for sampling precision and density, visibility and frustum occupancy. Unfortunately, positioning system error negatively impacts all of these requirements, with the severity generally being in the order of frustum occupancy, measurement variation and occlusion. Additionally, pose error imposes image-based registration constraints on view plan generation. In turn, this increases the length of the

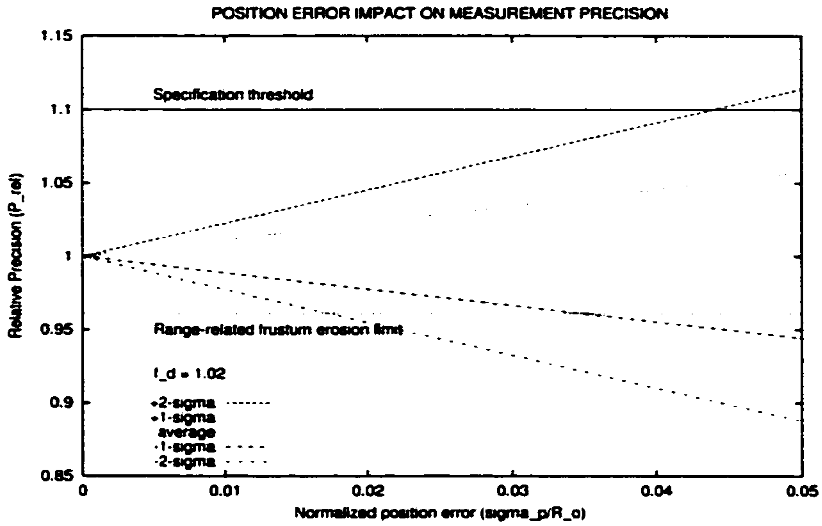


Figure 7.5: Pose Position Error Impact on Measurement Precision

view plan and the time span of all associated model reconstruction steps. Finally, collision avoidance planning is exacerbated by pose uncertainty.

Consequently, during view planning, we need to take both sensor and positioning system error models into account. It is clearly pointless to attempt subtle view planning optimization beyond the precision of the positioning system. This chapter has provided a qualitative and quantitative analysis of pose error effects on a common type of range camera. Based on this analysis, countermeasures have been identified to mitigate pose error effects on object reconstruction.

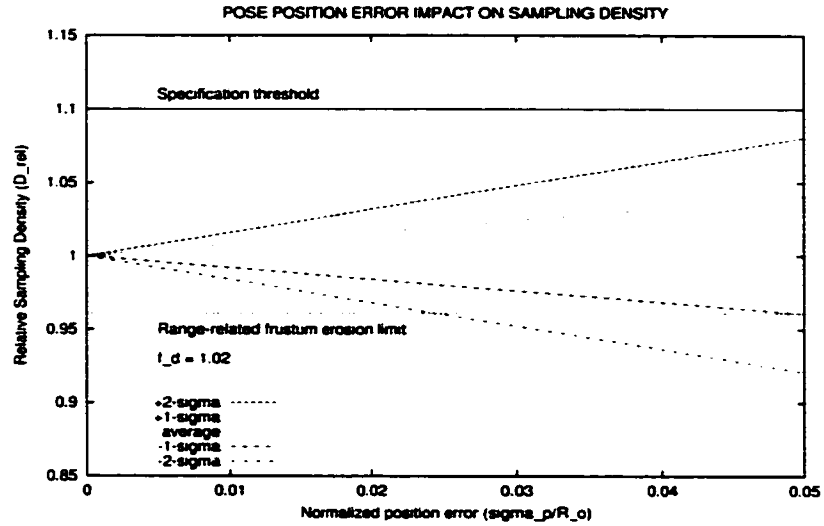


Figure 7.6: Pose Position Error Impact on Sampling Density

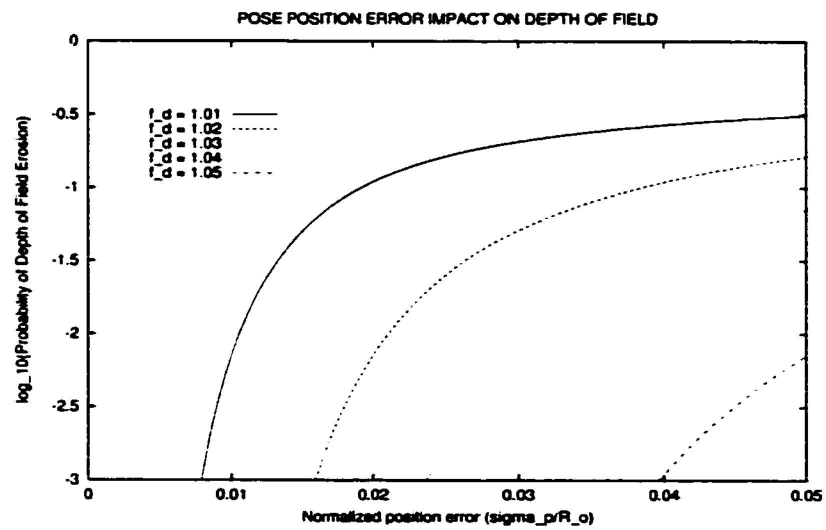


Figure 7.7: Pose Position Error Impact on Depth of Field

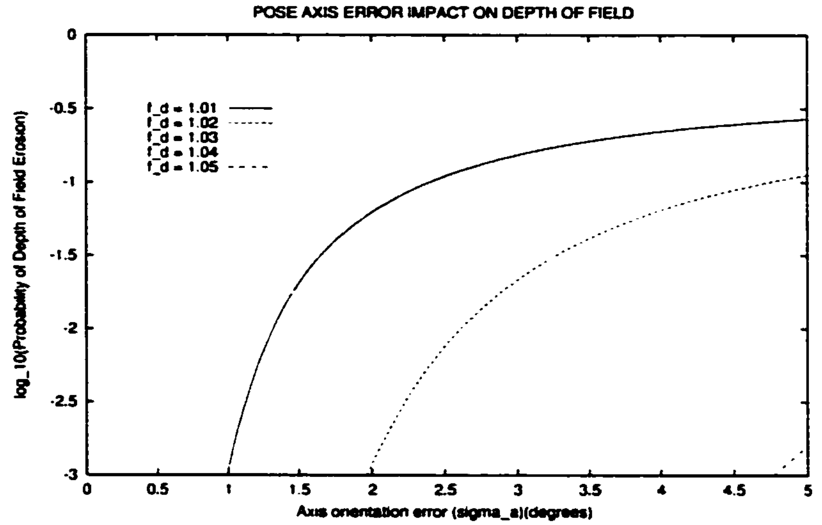


Figure 7.8: Axis Error Impact on Depth of Field

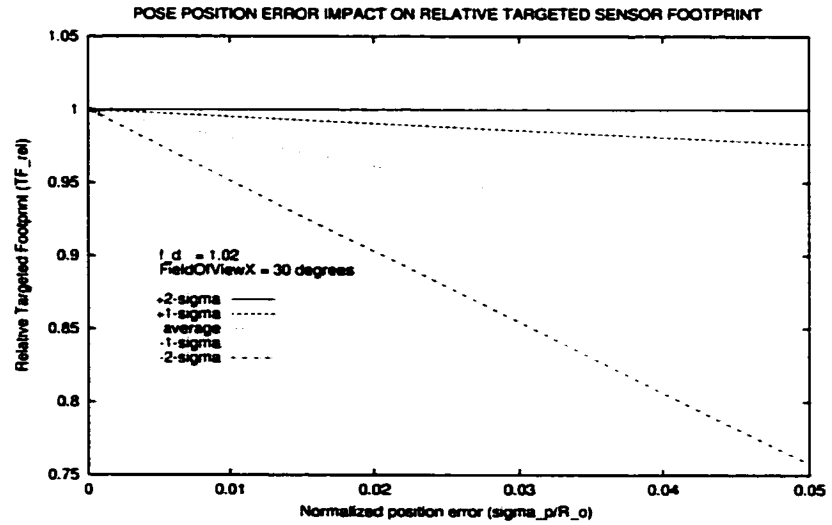


Figure 7.9: Pose Position Error Impact on Relative Targeted Sensor Footprint

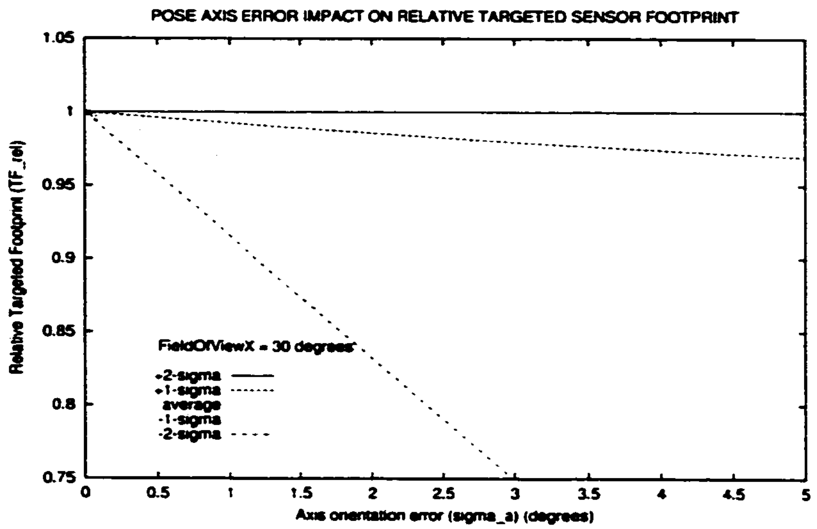


Figure 7.10: Pose Axis Error Impact on Relative Targeted Sensor Footprint

Measurement	**Pose Position Error**	**Pose Axis Error**	**Pose Twist Error**
Precision	$\mu P_{rel} = 1 + \frac{1}{3f_d^2}\frac{\sigma_p^2}{R_o^2}$ $\sigma P_{rel}^2 = \frac{4}{3f_d^2}\frac{\sigma_p^2}{R_o^2}$	$\mu P_{rel} = 1 - \frac{3\sigma_a^2}{4}$ $\sigma P_{rel}^2 = 0$	$\mu P_{rel} = 1$ $\sigma P_{rel}^2 = 0$
Sampling Density	$\mu D_{rel} = 1 + \frac{1}{3f_d^2}\frac{\sigma_p^2}{R_o^2}$ $\sigma D_{rel}^2 = \frac{2}{3f_d^2}\frac{\sigma_p^2}{R_o^2}$	$\mu D_{rel} = 1 - \frac{\sigma_a^2}{4}$ $\sigma D_{rel}^2 = 0$	$\mu D_{rel} = 1$ $\sigma D_{rel}^2 = 0$

Table 7.3: Single View Pose Error Measurability Effects: Line-scan Sensor

Error	**Field of View Erosion**
Position	$\mu_{TF_{rel}} = 1 - \frac{(\Phi_x T\Phi_2 + \Phi_x + 2T\Phi_2)}{2f_d\Phi_x T\Phi_2}\sqrt{\frac{2}{3\pi}}\frac{\sigma_p}{R_o} + \frac{(T\Phi_2 + 1)}{3\pi f_d{}^2\Phi_x T\Phi_2}\frac{\sigma_p{}^2}{R_o{}^2}$ ${\sigma_{TF_{rel}}}^2 = \frac{(\Phi_x T\Phi_2 + \Phi_x + 2T\Phi_2)^2}{4f_d{}^2\Phi_x{}^2T^2\Phi_2}\left(\frac{\pi - 2}{3\pi}\right)\frac{\sigma_p{}^2}{R_o{}^2}$
Axis	$\mu_{TF_{rel}} = 1 - \frac{(\Phi_x + 2T\Phi_2)}{\pi T\Phi_2\Phi_x}\sqrt{\frac{2}{\pi}}\sigma_a + \frac{\sigma_a{}^2}{2\pi T\Phi_2\Phi_x}$ ${\sigma_{TF_{rel}}}^2 = \frac{\sigma_a{}^2}{8\pi^3T^2\Phi_2\Phi_x{}^2}[(4T^2\Phi_2 + \Phi_x{}^2)(\pi^3 - 16) + 8T\Phi_2\Phi_x(\pi^2 - 8)]$
Twist	$\mu_{TF_{rel}} = 1 - \frac{\Phi_x{}^2 + 4T^2\Phi_2}{8\Phi_x T\Phi_2}\sqrt{\frac{2}{\pi}}\sigma_t + \frac{\sigma_t{}^2}{2}$ ${\sigma_{TF_{rel}}}^2 = \frac{(\Phi_x{}^2 + 4T^2\Phi_2)^2}{64\Phi_x{}^2T^2\Phi_2}\left(\frac{\pi - 2}{\pi}\right)\sigma_t{}^2$

Table 7.4: Single View Pose Error Field of View Erosion: Line-scan Sensor

Error	**Depth of Field Erosion**
Position	$P[p_z > (f_d - 1)R_o] = 1 - \Phi[\frac{\sqrt{3}(f_d - 1)}{(\sigma_p/R_o)}]$
Axis	$P[\alpha > \frac{(f_d - 1)}{f_d\Phi_2}] = 1 - \Phi[\frac{(f_d - 1)}{f_d\Phi_2\sigma_a}]$
Twist	nil

Table 7.5: Single View Pose Error Depth of Field Erosion: Line-scan Sensor

	Condition
1	Symmetric image: $N_x = N_y = N$
2	Scan length optimized for even sampling in x and y: $L_y = f_d R_o \Phi_x$
3	Optimal stand-off distance: $z = f_d R_o$
4	Measurements normal to the surface: $\theta_{xz} = \theta_{yz} = 0$
5	Measurements near sensor boresight: $\phi_{xz} = 0$, $x = y = 0$

Table 7.6: Optimal Scanning Conditions: Line-scan Sensor

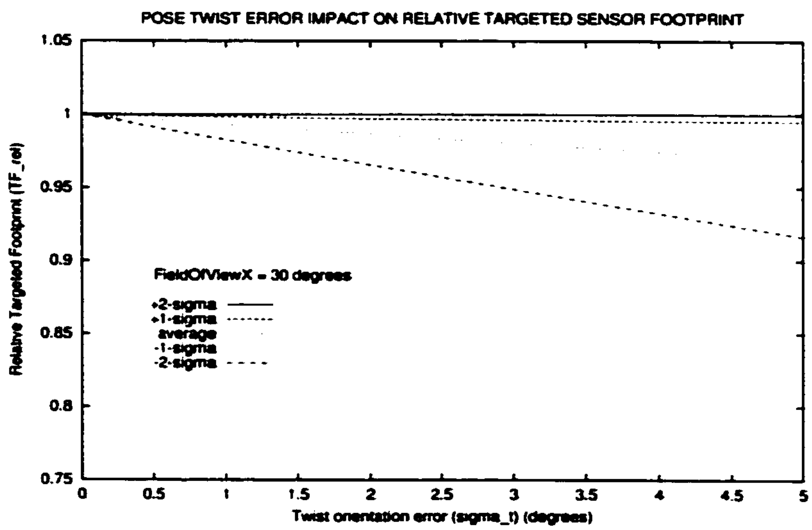


Figure 7.11: Pose Twist Error Impact on Relative Targeted Sensor Footprint

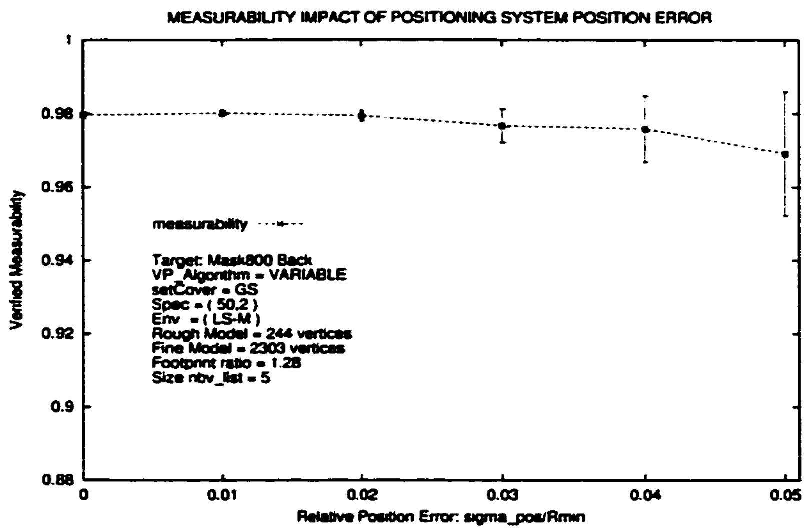


Figure 7.12: Multiple View Pose Error Effects - Position Error

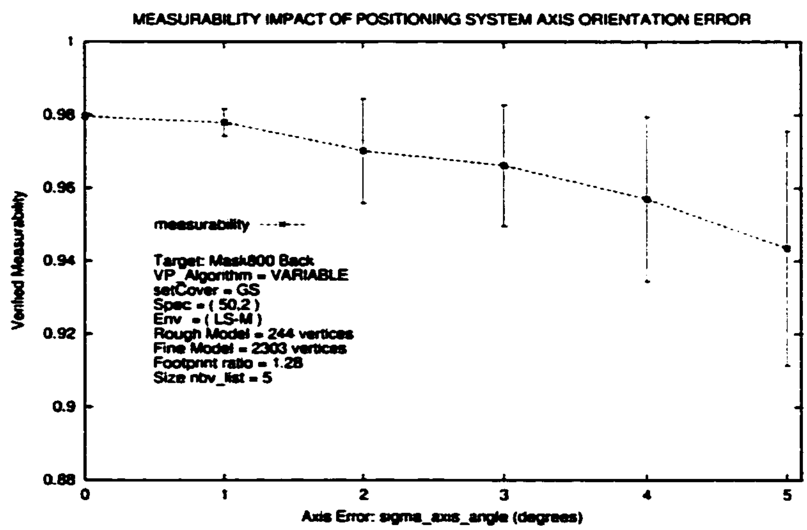


Figure 7.13: Multiple View Pose Error Effects - Axis Error

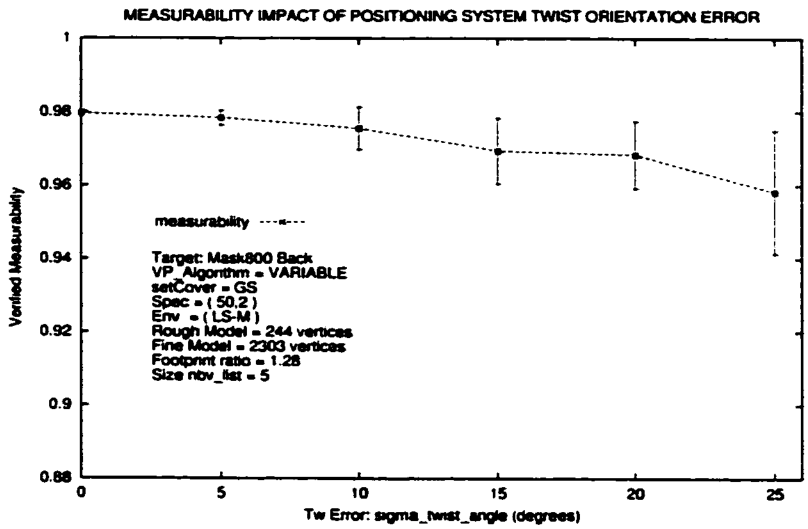


Figure 7.14: Multiple View Pose Error Effects - Twist Error

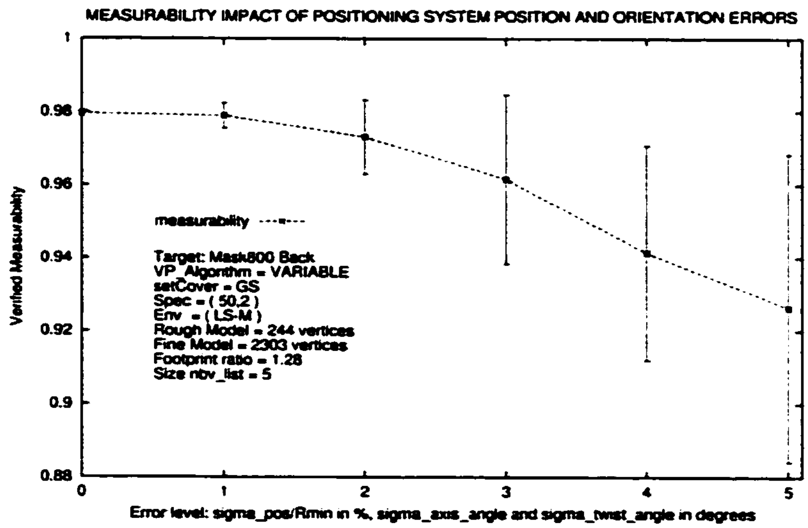


Figure 7.15: Multiple View Pose Error Effects - Combined Error

Chapter 8

Conclusion

This chapter summarizes the research, examines its capabilities and limitations and describes open view planning issues for future research.

8.1 Summary

This thesis presents and analyzes the performance of a model-based approach to view planning for automated object reconstruction. The view planning approach is "performance-oriented" in the sense that it is based on a set of explicit quality requirements expressed in a model specification and incorporates performance models for the range camera and positioning system. A new theoretical framework for the view planning problem is presented as an instance of the set covering problem with a registration constraint and is formulated as an integer programming problem. The theoretical framework facilitates a more intuitive understanding of the nature of view planning. It also provides an explicit mathematical formulation amenable to a well understood set of exact and approximate solution techniques. The measurability impact of pose error is studied in detail and counter-measures are presented to mitigate its effects. Methods for sparse sampling of object surface space and viewpoint space are presented and their performance is characterized.

In the introduction, we developed a detailed set of view planning requirements for automated object reconstruction and identified the main open issues as accuracy, robustness and efficiency. We believe the research reported here has made sufficient advances in all three areas to enable construction of a semi-automated performance-oriented object reconstruction system for the class of objects with benign reflectance characteristics. The steps remaining to turn the current experimental view planning algorithms into a complete prototype system involve mainly system development and integration issues rather than research. To expand this capability to the broader class of objects with wide reflectance variations will require research into view planning techniques addressing range camera dynamic range limitations.

The current work addresses the open view planning issues as follows:

- *Accuracy* is addressed by incorporating an explicit quantified specification of model quality objectives and by employing models of camera and positioning system performance.
- *Robustness* is addressed by the foregoing plus a robust, self-terminating set covering algorithm.
- *Efficiency* is addressed by multi-stage problem solving, sparse and optimal discretization of object surface space and viewpoint space and optional rough model segmentation. Additionally, the structure of the 3M algorithm is inherently parallelizable.

8.2 Capabilities and Limitations

This thesis addresses all view planning requirements for automated object reconstruction as identified in section 1.2 with the exception of some aspects

Category	**Requirement**	**Scott**
General	Model quality specification	Y
	Generalizable algorithm	Y
	Generalized viewpoints	Y
	View overlap	P
	Robust	P
	Efficient	Y
	Self-terminating	Y
Object	Minimal a priori knowledge	Y
	No Shape constraints	Y
	Material constraints	N
Sensor	Frustum modeled	Y
	Shadow effect	Y
	Measurement performance	Y
Positioning System	6D pose	Y
	Pose constraints	Y
	Positioning performance	Y

Table 8.1: Performance-Oriented View Planning Comparison

of the overlap requirement, material specular reflectance properties and some range camera artifacts. An evaluation is summarized at Table 8.1. The legend for the comparison table is as follows: requirement satisfied (Y), not satisfied (N) and partially satisfied (P). The work has limitations, however, and some issues remain open for a complete and general purpose solution to the view planning problem. We now discuss these capabilities and limitations, their significance and possible solutions.

Model Quality Specification Chapter 1 introduced the performance-oriented view planning concept. The modified measurability matrix (3M) view planning strategy incorporates a model specification for measurement precision and sampling density.

Generalizable Algorithm The 3M algorithm is generalizable to a variety of sensor and positioning system types. For the work presented here, two generic classes of range camera have been modeled - the "line-scan" and "random-access" configurations.

Generalized Viewpoints Generalized viewpoints are implemented. The imaging environment specification defines the system configuration, whether parameters are fixed or configurable, plus their nominal values and limits.

View Overlap The problem formulation includes a viewpoint correlation constraint which is equivalent to an image overlap constraint. Image overlap is a necessary but not sufficient condition for registration. The constraint needs to be extended to require sufficient shape complexity in the overlap region. The improved constraint should also take measurement uncertainty into account. While important for the purpose of completing the theoretical formulation, extending the constraint to cover shape complexity is not an urgent research issue. There is a strong probability that intrinsic redundancy in a view plan for a complex object will meet registration requirements as a by-product of the coverage and overlap constraints.

Robust The 3M algorithm partially satisfies the robustness requirement. Several features of the algorithm contribute to robustness i.e. a multi-stage problem solving approach, built-in sensor and positioning system performance models, a tolerance for surface sampling error in the rough model, counter-measures available to deal with pose error and a robust set covering algorithm. Measurement challenges remain, however, notably specular reflectance properties as well as some types of noise artifact.

Efficient The algorithm addresses the efficiency challenge by multi-stage problem solving, sparse and optimal discretization of object surface space and viewpoint space, and optional rough model segmentation. We have demonstrated good computational efficiency for representative reconstruction tasks on a platform that is slow by contemporary computing standards. The algorithm is readily parallelizable. High performance parallel computing clusters are now readily available and affordable. The performance and cost-effectiveness of computing platforms continue to improve steadily, making efficiency less of an issue. However, until a parallel version has been built and tested, we assess the capability of the 3M algorithm as a qualified "Y" with respect to the efficiency requirement.

Self-terminating We find a greedy search set covering algorithm modified with a registration constraint provides satisfactory results for most object reconstruction problems. The algorithm is fast, robust and self-terminating.

Limited a priori knowledge The only prior object knowledge required by the algorithm is approximate object bounding dimensions and centroid.

Shape constraints Experimentation has provided rules of thumb for rough model sampling rates and error tolerance. We anticipate that the greatest challenge presented by object shape complexity will be at the rough modeling rather than the fine modeling stage. Some user intervention may be required at the rough modeling stage for difficult shapes.

Material constraints The current work assumes objects characterized by Lambertian scattering. The implications of limited sensor dynamic range remain an important open issue. As has been previously noted, dynamic range problems manifest themselves as dropouts and outliers. Part of the solution will be found in sensor improvements. Better quality range cameras already incorporate automatic gain control in the laser-detector control loop. The improved effective dynamic range is beneficial but does not completely solve the problem. Further improvements should be possible by incorporating knowledge of object reflectance properties in the exploratory model. In the case of material uniformity, reflectance data could be gathered by a small number of point measurements by a separate instrument. For the more general case of non-uniform reflectance, it will be necessary to incorporate reflectance measurement and modeling as part of the scene exploration stage.

Frustum The sensor depth of field, field of view and scan length or scan arc are modeled.

Shadow effect Visibility is computed from both the laser source and receiver.

Measurement performance We assess the capability of the 3M algorithm as a qualified "Y" in this category. Measurability estimation error has been shown to be low and good measurability results are obtained at fairly coarse rough model sampling rates. The current work addresses the most important aspects of range camera measurement performance - in particular, spatial variation of precision and sampling density within the frustum and surface inclination effects. Other important measurement artifacts remain, notably those associated with geometric step edges, reflectance step edges and multiple reflections. Methods to deal with geometric step edges include "space-time" processing [CL96] and more sophisticated low level signal processing within the sensor. Object reflectance modeling would provide the additional information necessary to handle reflectance step edges. However, reflectance modeling is a highly complex process in itself [MTY01]. Multiple reflections result in measurement outliers. Computing multiple reflections is computationally burdensome but is necessary in more advanced forms of reflectance modeling. The extent to which it is appropriate and effective to integrate object reflectance modeling in view planning is to be determined.

6D pose The optimal scanning zone and decoupled algorithms provide a near-optimal sampling of 6D pose space unconstrained except for external restrictions applied by the user.

Pose constraints Pose constraints are applied at the viewpoint generation stage of the 3M algorithm.

Positioning system performance Pose error is modeled, allowing the user to select appropriate counter-measures.

8.3 Future Work

The capabilities and limitations discussed above point to research topics addressing the remaining open issues in view planning:

- *Accuracy and robustness*
  - Measuring and modeling object reflectance to handle shiny objects and compensate for limited range camera dynamic range.
  - Avoiding or compensating for geometric step edges, reflectance step edges and multiple reflections.

- *Efficiency*
  - Fine tuning discretization schemes in object surface space and viewpoint space.
  - Tuning the performance of approximate set covering algorithms to the data characteristics associated with the view planning problem.
  - Parallel implementation and visibility analysis speed up. (Development)
  - Automated surface segmentation for view planning purposes.
- *Theory*
  - Extend the theoretical framework to include shape complexity in the overlap constraint.
  - Develop a theoretical basis for optimal discretization of object surface space and viewpoint space.
- *General*
  - Develop performance benchmarks for objective comparison of view planning algorithm performance.

Appendix A

Symbols, Acronyms and Definitions

Symbol	**Definition**
3M	Modified measurability matrix
α	Pose axis error cone half-angle
$\vec{a}$	Sensor boresight axis
a_{kj}	Registration adjacency matrix element
A	Registration adjacency matrix $A = [a_{kj}]$
A_{rel}	Relative overlap area
b	Optical baseline
c	Number of segments in a segmented measurability matrix
c_{kj}	Cumulative registration adjacency matrix element
C	Cumulative registration adjacency matrix $C = [c_{kj}]$
C_x, C_y, C_z	Calibrated geometric noise coefficients in x-, y- and z-axes
CMM	Coordinate Measuring Machine
d_t	Number of twist quantization intervals
d_v	Viewpoint positional sub-sampling rate
D_{rel}	Relative sampling density (error vs. error-free)
DOF	Depth of field
D/W	Depth-to-width ratio
e_v	View plan efficiency $e_v = n/n_{Opt}$
EGI	Extended Gaussian Image
f_d	Sensor stand-off distance factor $f_d = 1 + \delta$, $\delta \ll 1$
f_o	Viewpoint positional over-sampling rate

Symbol	Definition
f_s	Sensor cross-section $f_s = \sqrt{F_s}$
F_p	Surface patch footprint (area)
F_s	Sensor footprint (area)
FOM	Figure of merit
FOV	Field of view
G_r	Viewpoint registration graph
GA	Genetic algorithm
GS	Greedy search
I	Imaging work space
IP	Integer programming
λ_s	Controllable sensor parameters
L_y	Sensor linear scan length along y-axis
LP	Linear programming
MVC	Mass Vector Chain
m_{ij}	Measurability matrix element
m_v	Verified measurability
$\mathbf{M}$	$\mathbf{M} = [m_{ij}]$
$\mathbf{M}_{i,V}$	Measurability matrix row vector
$\mathbf{M}_{S,j}$	Measurability matrix column vector
n	Size of NBV list
n_{Best}	Best size of view plan found for this problem
n_{Opt}	Optimum size of view plan
N	Next-best-view list or view plan
NBV	Next-best-view
N_x, N_y	Number of range image samples in x- and y-axes
ϕ_{xz}	Scan angle in sensor xz-plane
Φ_x	Sensor field of view in the scanning plane
Φ_2	$\Phi_2 = \Phi_x/2$
$\Phi[z]$	Normal distribution function $\Phi[z] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-u^2/2}\,du$
p_x, p_y, p_z	Pose position error in x-, y- and z-axes
$\vec{p}$	Sensor pose position
P_{rel}	Relative measurement precision (error vs. error-free)
r_f	Surface patch's footprint ratio $r_f = F_p/F_s$
ρ_{rel}	Relative sampling density
ρ_v	Positional density of viewpoints

Symbol	**Definition**
$\hat{\rho}_z$	Estimated sampling density
R	Stand-off distance
R_a	Stand-off adjustment range
R_o	Optimum sensor scanning range
R_{min}	Minimum sensor scanning range
R_{max}	Maximum sensor scanning range
R_{xz}	Slant range in x-z plane
σ_a	Pose axis angle uncertainty standard deviation
σ_p	Pose position uncertainty standard deviation
σ_t	Pose twist angle uncertainty standard deviation
σ_{kj}	Cross-correlation of viewpoints v_k and v_j
σ_{rm}	Rough model surface noise standard deviation
$\hat{\sigma}_x$, $\hat{\sigma}_y$, $\hat{\sigma}_z$	Estimated precision in x-, y- and z-axes
σ_v	Viewpoint correlation
Σ	Cross-correlation matrix $\Sigma = [\sigma_{kj}]$
s	Size of object surface space
S	Object surface space
$\mathbf{S_o}$	Optimal viewpoint zone
SA	Simulated annealing
SCP	Set covering problem
$S\theta$,$C\theta$,$T\theta$	$\sin\theta, \cos\theta, \tan\theta$
θ_{xz}, θ_{yz}	Laser scanning ray incidence angles in xz- and yz-planes
t_r	Registration threshold
t_w	Sensor twist angle (rotation about the boresight)
t_{xz}, t_{yz}	Grazing angle thresholds in x-z and y-z planes
TF	Targeted footprint
TF_{rel}	Relative targeted footprint (error vs. error-free)
$U_i(\theta)$	Inverse unit step function, $(U_i(\theta) = 1 - U(\theta))$
$\mathbf{v}$	Sensor pose $\mathbf{v} =$ (position $\vec{p}$, axis $\vec{a}$, twist t_w)
v	Size of viewpoint space
V	Viewpoint space
$V\theta$	$vers\,\theta = 1 - \cos\theta$
VPP	View planning problem
x_j	Binary viewpoint variable
$\mathbf{X}$	Vector of binary viewpoint variables $\mathbf{X} = [x_j]$

Appendix B

Registration Constraint Example

This appendix illustrates the registration constraint developed in Chapter 4 with a simple example.

Consider the set of four candidate viewpoints shown with their global registration graph in Figure B.1. The base registration adjacency matrix A can be written[1] as

$$\begin{bmatrix} 0 & 0 & x_1x_3 & 0 \\ 0 & 0 & 0 & x_2x_4 \\ x_1x_3 & 0 & 0 & x_3x_4 \\ 0 & x_2x_4 & x_3x_4 & 0 \end{bmatrix}$$

from which we obtain the cumulative registration adjacency matrix $C = A^1 + A^2 + A^3$

[1] As we are interested only in simple connectivity, we take the liberty of using boolean algebra in matrix manipulation and simplification.

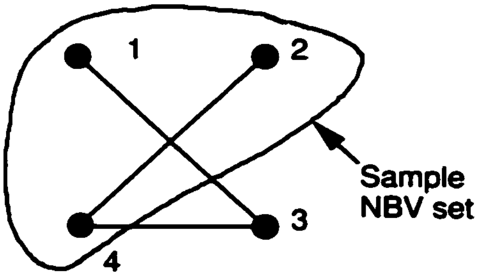


Figure B.1: Example Registration Graph G_r and NBV Set

$$\begin{bmatrix} x_1x_3 & x_1x_2x_3x_4 & x_1x_3 & x_1x_3x_4 \\ x_1x_2x_3x_4 & x_2x_4 & x_2x_3x_4 & x_2x_4 \\ x_1x_3 & x_2x_3x_4 & x_3(x_1+x_4) & x_3x_4 \\ x_1x_3x_4 & x_2x_4 & x_3x_4 & x_4(x_2+x_3) \end{bmatrix}$$

or simply as the following where "-" signifies "don't care".

$$\begin{bmatrix} - & x_1x_2x_3x_4 & x_1x_3 & x_1x_3x_4 \\ - & - & x_2x_3x_4 & x_2x_4 \\ - & - & - & x_3x_4 \\ - & - & - & - \end{bmatrix}$$

Using equation 4.6 and removing redundancies, the foregoing simplifies to the following global registration constraints which can be validated by inspection of the viewpoint registration graph G_r:

$$\begin{aligned} x_1x_2x_3x_4 &\geq x_1x_2, \\ x_1x_3x_4 &\geq x_1x_4, \\ x_2x_3x_4 &\geq x_2x_3. \end{aligned} \tag{B.1}$$

The constraints at Equation B.1 can be used to test whether an arbitrary viewpoint set meets the registration constraint. With the example NBV set as shown at Figure B.1, viewpoint variables x_1, x_2 and x_4 are true while viewpoint x_3 is false. Testing the registration graph in Figure B.1 against the constraints at equation set B.1 for this viewpoint set, we observe that the example candidate NBV set fails the registration test.

Appendix C

View Planning Impact of Pose Error

This appendix analyzes the effects of each component of pose error on each quality factor as illustrated in Table 7.1, beginning with pose position error. The analysis examines pose error effects on a single planned view. Results are summarized in Chapter 7. Key symbols and definitions are summarized at Appendix A.

C.1 Position Error

As a consequence of changing the location of the entire optical baseline, position error potentially impacts all model specification factors: frustum occupancy, visibility, measurement precision and sampling density.

C.1.1 Measurement Precision Impact

This section considers the effect of pose position error on measurement precision estimation.

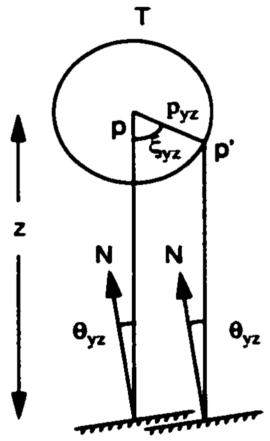


Figure C.1: Scanning Geometry Y-Z Plane

Precision Estimation Model From Equation 7.4, our model of relative measurement precision is

$$P_{rel} = \frac{z^2}{z'^2}\frac{\cos\theta'_{yz}}{\cos\theta_{yz}}. \tag{C.1}$$

Assume the viewpoint is optimally positioned at a stand-off distance slightly beyond the sensor minimum range i.e. $z = f_d R_o$, $f_d = 1 + \delta$, where $\delta \ll 1$. Sensor position is corrupted by position error $\vec{p} = (p_x, p_y, p_z)$. Then, as shown in Figure 3.7, z-axis range to the surface element is perturbed to $z' = z + p_z$, that is $z' = f_d R_o + p_z$.

With a line-scan camera subject to pose position error, the inclination angle in the y-z plane with respect to the laser source is unchanged - that is, $\theta_{yz} = \theta_{yz}'$. The scanning geometry is amplified at Figure C.1 for clarity. Consequently, P_{rel} simply becomes z'^2/z^2 from which we get

$$\begin{aligned} P_{rel} &= \frac{(f_d R_o + p_z)^2}{(f_d R_o)^2} \\ &= (1 + \frac{1}{f_d}(\frac{p_z}{R_o}))^2 \\ &= 1 + \frac{2}{f_d}(\frac{p_z}{R_o}) + \frac{1}{f_d^2}(\frac{p_z}{R_o})^2. \end{aligned} \tag{C.2}$$

We wish to estimate the statistics of the relative precision P_{rel} which is a function of the random variable p_z. From our positioning system error model, the expected value and variance of p_z are as follows:

$$\begin{aligned} E[p_z] &= 0, \\ E[p_z^2] &= \sigma_{p_z}^2 \\ &= \sigma_p^2/3. \end{aligned} \tag{C.3}$$

Statistics of P_{rel} Consequently,

$$\begin{aligned} E[P_{rel}] &= \mu_{P_{rel}} \\ &= 1 + \frac{2}{f_d R_o} E[p_z] + \frac{1}{f_d^2 R_o^2} E[p_z^2] \\ &= 1 + \frac{\sigma_p^2}{3 f_d^2 R_o^2}. \end{aligned} \tag{C.4}$$

To compute the variance of P_{rel},

$$\begin{aligned} E[(P_{rel} - \mu_{P_{rel}})^2] &= \sigma_{P_{rel}}^2 \\ &= E[P_{rel}^2] - (\mu_{P_{rel}})^2. \end{aligned} \tag{C.5}$$

Then,

$$E[P_{rel}^2] = E[(1 + \tfrac{2}{f_d}(\tfrac{p_z}{R_o}) + \tfrac{1}{f_d^2}(\tfrac{p_z}{R_o})^2)^2] \tag{C.6}$$

and, discarding terms over quadratic,

$$\begin{aligned} E[P_{rel}^2] &\approx E[1 + \tfrac{4}{f_d}\tfrac{p_z}{R_o} + \tfrac{6}{f_d^2}\tfrac{p_z^2}{R_o^2}] \\ &= 1 + \frac{2\sigma_p^2}{f_d^2 R_o^2}. \end{aligned} \tag{C.7}$$

So, substituting Equations C.7 and C.2 in Equation C.5, we have

$$\sigma_{P_{rel}}{}^2 = \frac{4}{3f_d^2}\frac{\sigma_p^2}{R_o^2}. \tag{C.8}$$

Therefore, the relative impact of the position component of pose error on measurement precision estimation is approximately

$$\mu_{P_{rel}} = 1 + \frac{1}{3f_d^2}\frac{\sigma_p^2}{R_o^2}, \tag{C.9}$$

$$\sigma_{P_{rel}} = \frac{2}{\sqrt{3}f_d}\frac{\sigma_p}{R_o}. \tag{C.10}$$

Precision Impact Interpretation In Figure 7.5, relative precision P_{rel} computed from Equations C.9 and C.10 is plotted against normalized position error σ_p/R_o. The latter will typically be very small i.e. $\sigma_p/R_o \ll 1$. Figure 7.5 can be interpreted as follows. One- and two-sigma curves for relative precision are displayed bracketing the nominal value of 1.0. Precision improves for values $P_{rel} < 1$ and deteriorates for values $P_{rel} > 1$. Any vertical slice through these curves defines the shape of the probability density function for relative precision for a given normalized position error. Horizontal lines at $P_{rel} = 1.1$ and $P_{rel} = 0.961$ indicate boundary conditions beyond which measurements will be rejected due to pose position error. The upper limit represents an example of a specified measurement precision limit from the model specification. It has been arbitrarily set at 1.1 in Figure 7.5 for illustration purposes. The lower limit, easily shown to be $1/{f_d}^2$, defines the point at which measurements are rejected as falling outside the sensor near-range limit. The limit is illustrated here for a value of $f_d = 1.02$, meaning viewpoints are optimized for a stand-off distance of 2% beyond the optimal sensor range R_o. The lower limit cut-off is one manifestation of frustum erosion, a topic treated later in more detail.

Examining this error component in isolation, it is apparent from the foregoing that pose position error has a low to moderate impact on measurement precision. The specification clipping effect can be mitigated by selecting suitably conservative model specification limits relative to sensor and positioning system capabilities. Means to mitigate frustum erosion effects will be addressed later.

C.1.2 Sampling Density Impact

Sampling Density Estimation Model This section considers the effect of pose position error on sampling density estimation. From Equation 7.9, our model for sampling density estimation is

$$D_{rel} = \frac{C^2\theta'_{xz}C^2\theta'_{yz}}{C^2\theta_{xz}C^2\theta_{yz}} \frac{{R_{xz}}^2C^2\theta_{yz} + {f_d}^2{R_o}^2C^2\theta_{xz}}{{R'_{xz}}^2C^2\theta'_{yz} + {f_d}^2{R_o}^2C^2\theta'_{xz}}. \tag{C.11}$$

However, as $\theta'_{yz} = \theta_{yz}$, Equation C.11 reduces to

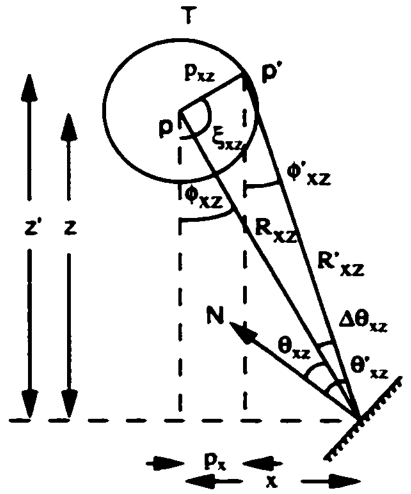


Figure C.2: Scanning Geometry X-Z Plane

$$D_{rel} = \frac{C^2\theta'_{xz}}{C^2\theta_{xz}} \frac{R_{xz}{}^2 C^2\theta_{yz} + f_d{}^2 R_o{}^2 C^2\theta_{xz}}{R'_{xz}{}^2 C^2\theta_{yz} + f_d{}^2 R_o{}^2 C^2\theta'_{xz}}. \tag{C.12}$$

Key Geometric Parameters We now wish to compute the statistics of D_{rel} which is a function of the random variables p_{xz} and ξ_{xz}, the magnitude and angle of position error in the scanning plane, respectively. It is useful to first compute simple expressions for the key geometric parameters. From Figure C.2, we note that

$$R_{xz}{}^2 = f_d{}^2 R_o{}^2 + x^2 \tag{C.13}$$

$$R'_{xz}{}^2 = (f_d R_o + p_z)^2 + (x - p_x)^2 \tag{C.14}$$

and

$$\frac{p_{xz}}{R_{xz}} = \frac{S\delta\theta_{xz}}{S[(\xi_{xz} - \phi_{xz}) + \delta\theta_{xz}]}. \tag{C.15}$$

Defining $r = p_{xz}/R_{xz} = (p_{xz}C\phi_{xz})/(f_d R_o)$ and setting $\psi_{xz} = \xi_{xz} - \phi_{xz}$, we have the relationship

$$\cot\delta\theta_{xz} = \frac{1 - rC\psi_{xz}}{rS\psi_{xz}} \tag{C.16}$$

from which we can determine

$$S\delta\theta_{xz} = \frac{rS\psi_{xz}}{\sqrt{1 + r^2 - 2rC\psi_{xz}}}, \tag{C.17}$$

$$C\delta\theta_{xz} = \frac{1 - rC\psi_{xz}}{\sqrt{1 + r^2 - 2rC\psi_{xz}}}. \tag{C.18}$$

Noting that $r \ll 1$, the foregoing expressions for $S\delta\theta_{xz}$ and $C\delta\theta_{xz}$ can be written as follows, where we ignore terms in r above quadratic

$$S\delta\theta_{xz} \approx rS\psi_{xz} + r^2 S\psi_{xz}C\psi_{xz}, \tag{C.19}$$

$$C\delta\theta_{xz} \approx 1 + \frac{r^2}{2}(C^2\psi_{xz} - 1). \tag{C.20}$$

As an additional precursor to computing the statistics of D_{rel}, we next compute the statistics of ξ_{xz}, r and $\delta\theta_{xz}$. We again only consider terms up to quadratic in r.

Statistics of ξ_{xz}

$$E[S\psi_{xz}] = E[C\psi_{xz}] = E[S\psi_{xz}C\psi_{xz}] = 0 \tag{C.21}$$

$$E[S^2\psi_{xz}] = E[C^2\psi_{xz}] = 1/2 \tag{C.22}$$

Statistics of r

$$E[r] = \frac{C\phi_{xz}}{f_d R_o} E[p_{xz}] = 0 \tag{C.23}$$

$$E[r^2] = \frac{C^2\phi_{xz}}{{f_d}^2 {R_o}^2} E[{p_{xz}}^2] = \frac{2C^2\phi_{xz}}{3{f_d}^2} \frac{{\sigma_p}^2}{{R_o}^2} \tag{C.24}$$

Statistics of $\delta\theta_{xz}$

$$E[S\delta\theta_{xz}] = E[rS\psi_{xz} + r^2 S\psi_{xz} C\psi_{xz}] = 0 \tag{C.25}$$

$$E[C\delta\theta_{xz}] = E[1 + \tfrac{r^2}{2}(C^2\psi_{xz} - 1)] = 1 - \frac{C^2\phi_{xz}}{6{f_d}^2} \frac{{\sigma_p}^2}{{R_o}^2} \tag{C.26}$$

$$E[S^2\delta\theta_{xz}] = E[r^2 S^2\psi_{xz}] = \frac{C^2\phi_{xz}}{3{f_d}^2} \frac{{\sigma_p}^2}{{R_o}^2} \tag{C.27}$$

$$E[C^2\delta\theta_{xz}] = E[1 + r^2(C^2\psi_{xz} - 1)] = 1 - \frac{C^2\phi_{xz}}{3{f_d}^2} \frac{{\sigma_p}^2}{{R_o}^2} \tag{C.28}$$

$$E[S\delta\theta_{xz} C\delta\theta_{xz}] = E[rS\psi_{xz} + r^2 S\psi_{xz} C\psi_{xz}] = 0 \tag{C.29}$$

Simplification of D_{rel} We now have most of the ingredients to calculate the statistics of the relative sampling density D_{rel}. First, it is convenient to rewrite Equation C.12 as the product of a constant term T1 and a variable term T2.

$$D_{rel} = \underbrace{\frac{\frac{{R_{xz}}^2}{{f_d}^2{R_o}^2}C^2\theta_{yz} + C^2\theta_{xz}}{C^2\theta_{xz}}}_{T1} \underbrace{\frac{C^2\theta'_{xz}}{\frac{(R'_{xz})^2}{{f_d}^2{R_o}^2}C^2\theta_{yz} + C^2\theta'_{xz}}}_{T2} \tag{C.30}$$

Consider T1. As $R_{xz}/(f_d R_o) = 1/C\phi_{xz}$, T1 reduces to

$$T1 = \frac{C^2\theta_{yz} + C^2\phi_{xz}C^2\theta_{xz}}{C^2\phi_{xz}C^2\theta_{xz}}. \tag{C.31}$$

Let $T2 = N2/D2$. Then,

$$D2 = \frac{{R'_{xz}}^2}{{f_d}^2{R_o}^2}C^2\theta_{yz} + C^2\theta'_{xz}. \tag{C.32}$$

Considering the first term in Equation C.32 and using Equation C.14, we have

$$\begin{aligned}\frac{{R'_{xz}}^2}{{f_d}^2{R_o}^2} &= 1 + \frac{2}{f_d}\frac{p_z}{R_o} + \frac{1}{{f_d}^2}\frac{{p_z}^2}{{R_o}^2} + T^2\phi_{xz} - \frac{2T\phi_{xz}}{f_d}\frac{p_x}{R_o} + \frac{1}{{f_d}^2}\frac{{p_x}^2}{{R_o}^2} \\ &= \frac{1}{C^2\phi_{xz}}(1 + t_r) \end{aligned} \tag{C.33}$$

where t_r is

$$t_r = \frac{2C^2\phi_{xz}}{f_d}\frac{p_z}{R_o} + \frac{C^2\phi_{xz}}{{f_d}^2}\frac{{p_z}^2}{{R_o}^2} - \frac{2S\phi_{xz}C\phi_{xz}}{f_d}\frac{p_x}{R_o} + \frac{C^2\phi_{xz}}{{f_d}^2}\frac{{p_x}^2}{{R_o}^2}. \tag{C.34}$$

Considering the term $C^2\theta'_{xz}$ in Equation C.32,

$$\begin{aligned} C^2\theta'_{xz} &= C^2(\theta_{xz} + \delta\theta_{xz}) \\ &= (C\theta_{xz}C\delta\theta_{xz} - S\theta_{xz}S\delta\theta_{xz})^2 \\ &= C^2\theta_{xz}C^2\delta\theta_{xz} - 2S\theta_{xz}C\theta_{xz}S\delta\theta_{xz}C\delta\theta_{xz} + S^2\theta_{xz}S^2\delta\theta_{xz}. \end{aligned} \tag{C.35}$$

Then, using the relationships from Equations C.19 and C.20,

$$S^2\delta\theta_{xz} \approx r^2 S^2\psi_{xz} \tag{C.36}$$

$$C^2\delta\theta_{xz} \approx 1 + r^2(C^2\psi_{xz} - 1) \tag{C.37}$$

$$S\delta\theta_{xz}C\delta\theta_{xz} \approx rS\psi_{xz} + r^2 S\psi_{xz}C\psi_{xz} \tag{C.38}$$

so that we have

$$\begin{aligned} C^2\theta'_{xz} &\approx C^2\theta_{xz}[1 + r^2(C^2\psi_{xz} - 1)] \\ &\quad -2S\theta_{xz}C\theta_{xz}[rS\psi_{xz} + r^2 S\psi_{xz}C\psi_{xz}] + S^2\theta_{xz}r^2S^2\psi_{xz} \\ &= C^2\theta_{xz}[1 + t_\theta] \end{aligned} \tag{C.39}$$

where

$$\begin{aligned} t_\theta &= -2rT\theta_{xz}S\psi_{xz} + r^2(C^2\psi_{xz} - 1) \\ &\quad -2r^2T\theta_{xz}S\psi_{xz}C\psi_{xz} + r^2T^2\theta_{xz}S^2\psi_{xz}. \end{aligned} \tag{C.40}$$

Now, combining the expressions for ${R'_{xz}}^2/({f_d}^2{R_o}^2)$ and $C^2\theta'_{xz}$ with Equation C.32, we have

$$\begin{aligned}
D2 &= \frac{{R'_{xz}}^2}{{f_d}^2{R_o}^2}C^2\theta_{yz} + C^2\theta'_{xz} \\
&= \frac{C^2\theta_{yz}}{C^2\phi_{xz}}[1+t_r] + C^2\theta_{xz}[1+t_\theta] \\
&= \frac{C^2\theta_{yz} + C^2\theta_{xz}C^2\phi_{xz}}{C^2\phi_{xz}}[1 + \frac{C^2\theta_{yz}}{C^2\theta_{yz} + C^2\theta_{xz}C^2\phi_{xz}}t_r + \frac{C^2\theta_{xz}C^2\phi_{xz}}{C^2\theta_{yz} + C^2\theta_{xz}C^2\phi_{xz}}t_\theta] \\
&= \frac{C^2\theta_{yz} + C^2\theta_{xz}C^2\phi_{xz}}{C^2\phi_{xz}}[1+t] \qquad \text{(C.41)}
\end{aligned}$$

where

$$t = \frac{C^2\theta_{yz}}{C^2\theta_{yz} + C^2\theta_{xz}C^2\phi_{xz}}t_r + \frac{C^2\theta_{xz}C^2\phi_{xz}}{C^2\theta_{yz} + C^2\theta_{xz}C^2\phi_{xz}}t_\theta. \qquad \text{(C.42)}$$

Now, the expression for T2 from Equation C.32 can be simplified to

$$T2 = \frac{C^2\theta_{xz}C^2\phi_{xz}}{C^2\theta_{yz} + C^2\theta_{xz}C^2\phi_{xz}}\frac{[1+t_\theta]}{[1+t]}. \qquad \text{(C.43)}$$

Statistics of D_{rel} Finally, combining Equations C.31 and C.43, we have the following expression for D_{rel}:

$$D_{rel} = \frac{[1+t_\theta]}{[1+t]}. \qquad \text{(C.44)}$$

As $|t| < 1$, the denominator of Equation C.44 can be expressed as a series expansion. After expansion and collecting terms up to quadratic in t_θ and t, we have the following second order approximations for D_{rel} and ${D_{rel}}^2$:

$$D_{rel} \approx 1 - t + t^2 + t_\theta - tt_\theta, \tag{C.45}$$

$$D_{rel}{}^2 \approx 1 - 2t + 3t^2 + 2t_\theta - 4tt_\theta + t_\theta{}^2. \tag{C.46}$$

Using the interim results previously developed, the statistics of terms in t_θ and t can be expressed as follows.

$$E[t_\theta] = \frac{C^2\phi_{xz}}{3f_d{}^2}\frac{\sigma_p{}^2}{R_o{}^2}(T^2\theta_{xz} - 1) \tag{C.47}$$

$$E[t_\theta{}^2] = \frac{4T^2\theta_{xz}C^2\phi_{xz}}{3f_d{}^2}\frac{\sigma_p{}^2}{R_o{}^2} \tag{C.48}$$

$$E[t_r] = \frac{2C^2\phi_{xz}}{3f_d{}^2}\frac{\sigma_p{}^2}{R_o{}^2} \tag{C.49}$$

$$E[t_r{}^2] = \frac{8C^2\phi_{xz}}{3f_d{}^2}\frac{\sigma_p{}^2}{R_o{}^2} \tag{C.50}$$

$$E[t] = \frac{C^2\phi_{xz}}{3f_d{}^2}\frac{\sigma_p{}^2}{R_o{}^2}\frac{[2C^2\theta_{yz} + C^2\phi_{xz}(S^2\theta_{xz} - C^2\theta_{xz})]}{C^2\theta_{yz} + C^2\theta_{xz}C^2\phi_{xz}} \tag{C.51}$$

$$E[t^2] = \frac{4C^2\phi_{xz}}{3f_d{}^2}\frac{\sigma_p{}^2}{R_o{}^2}\frac{[2C^4\theta_{yz} + C^4\phi_{xz}S^2\theta_{xz}C^2\theta_{xz}]}{(C^2\theta_{yz} + C^2\theta_{xz}C^2\phi_{xz})^2} \tag{C.52}$$

$$E[t_\theta t_r] = 0 \tag{C.53}$$

$$E[tt_\theta] = \frac{4S^2\theta_{xz}C^4\phi_{xz}}{3f_d{}^2}\frac{\sigma_p{}^2}{R_o{}^2}\frac{1}{[C^2\theta_{yz} + C^2\theta_{xz}C^2\phi_{xz}]} \tag{C.54}$$

After these laborious calculations, we can finally express the statistics of D_{rel} from Equations C.45 and C.46 as follows after collecting terms up to quadratic in σ_p/R_o and simplifying:

$$\begin{aligned} E[D_{rel}] &= \mu_{D_{rel}} \\ &= 1 + \frac{C^2\theta_{yz}C^2\phi_{xz}}{3C^2\theta_{xz}f_d{}^2}\frac{\sigma_p{}^2}{R_o{}^2}\frac{4C^2\theta_{xz}C^2\theta_{yz} - 3C^2\theta_{xz}C^2\phi_{xz} + C^2\theta_{yz}}{(C^2\theta_{yz} + C^2\theta_{xz}C^2\phi_{xz})^2} \end{aligned} \tag{C.55}$$

$$E[D_{rel}{}^2] = 1 + \frac{2C^2\theta_{yz}C^2\phi_{xz}}{C^2\theta_{xz}f_d{}^2}\frac{\sigma_p{}^2}{R_o{}^2}\frac{2C^2\theta_{xz}C^2\theta_{yz} - C^2\theta_{xz}C^2\phi_{xz} + C^2\theta_{yz}}{(C^2\theta_{yz} + C^2\theta_{xz}C^2\phi_{xz})^2} \tag{C.56}$$

so the variance and standard deviation of D_{rel} are

$$\sigma_{D_{rel}}{}^2 = \frac{4C^4\theta_{yz}C^2\phi_{xz}(C^2\phi_{xz}+1)}{3C^2\theta_{xz}f_d{}^2(C^2\theta_{yz} + C^2\theta_{xz}C^2\phi_{xz})^2}\frac{\sigma_p{}^2}{R_o{}^2}, \tag{C.57}$$

$$\sigma_{D_{rel}} = \frac{2C^2\theta_{yz}C\phi_{xz}\sqrt{C^2\phi_{xz}+1}}{\sqrt{3}C\theta_{xz}f_d(C^2\theta_{yz} + C^2\theta_{xz}C^2\phi_{xz})}\frac{\sigma_p}{R_o}. \tag{C.58}$$

In summary, the relative impact of pose position error on sampling density estimation is approximately

$$\mu_{D_{rel}} = 1 + \frac{C^2\theta_{yz}C^2\phi_{xz}}{3C^2\theta_{xz}f_d{}^2}\frac{\sigma_p{}^2}{R_o{}^2}\frac{4C^2\theta_{xz}C^2\theta_{yz} - 3C^2\theta_{xz}C^2\phi_{xz} + C^2\theta_{yz}}{(C^2\theta_{yz} + C^2\theta_{xz}C^2\phi_{xz})^2}, \tag{C.59}$$

$$\sigma_{D_{rel}} = \frac{2C^2\theta_{yz}C\phi_{xz}\sqrt{C^2\phi_{xz}+1}}{\sqrt{3}C\theta_{xz}f_d(C^2\theta_{yz}+C^2\theta_{xz}C^2\phi_{xz})}\frac{\sigma_p}{R_o}. \tag{C.60}$$

For an optimum scanning geometry such that $\phi_{xz} = \theta_{xz} = \theta_{yz} = 0$, the statistics of relative sampling density D_{rel} reduce to the following:

$$\mu_{D_{rel}} = 1 + \frac{1}{3{f_d}^2}\frac{{\sigma_p}^2}{{R_o}^2}, \tag{C.61}$$

$$\sigma_{D_{rel}} = \frac{\sqrt{2}}{\sqrt{3}f_d}\frac{\sigma_p}{R_o}. \tag{C.62}$$

Sampling Density Impact Interpretation The impact of pose position error on sampling density estimation calculated at Equations C.61 and C.62 is shown in Figure 7.6 for the optimum scanning geometry. Sampling density improves for values $D_{rel} > 1$ and deteriorates for values $D_{rel} < 1$. The effects for sampling density will be seen to be very close to those for measurement precision as shown at Figure 7.5 and can be interpreted in a similar manner. As was the case for measurement precision, it is apparent from the foregoing that pose position error has a low to moderate impact on sampling density.

C.1.3 Visibility Impact

Visibility or occlusion effects depend on object shape, optical baseline length and sensor-object relative pose. The optical baseline is fixed and camera pose has been optimized in the viewpoint generation process. Relative visibility effects therefore depend mainly on object geometry. They cannot be quantified in the absence of a specific object shape. However, from the previous analysis, we can observe that the visibility impact of pose position error will be nil to low. This is due to the very small changes in observation geometry between surface points, laser source and optical receiver as a consequence of the low normalized position error σ_p/R_o anticipated for most imaging environments.

C.1.4 Frustum Occupancy Impact

This section considers the effect of pose position error on frustum occupancy.

Frustum Erosion From Equation 7.10, the targeted footprint is

$$TF = \underbrace{(2f_dR_oT\Phi_2)}_{width}\underbrace{(L_y)}_{length}. \tag{C.63}$$

Pose error $\vec{p} = (p_x, p_y, p_z)$ erodes coverage as follows. p_x reduces the width by $|p_x|$, p_y reduces the length by $|p_y|$ while p_z reduces the width by $T\Phi_2|p_z|$. The later relationship can be seen from the following. For $p_z < 0$, width is reduced by $2T\Phi_2|p_z|$. For $p_z > 0$, width is increased by $2T\Phi_2|p_z|$. However, the increase for $p_z > 0$ is of no benefit to the planning of a single view. The net effect of z-axis pose error is a coverage reduction of $T\Phi_2|p_z|$. Consequently, the portion of the targeted footprint covered by the viewpoint subject to pose position error is

$$TF' = (2f_dR_oT\Phi_2 - T\Phi_2|p_z| - |p_x|)(L_y - |p_y|). \tag{C.64}$$

Then, TF_{rel} is

$$TF_{rel} = \frac{TF'}{TF}$$

$$= \frac{(2f_dR_oT\Phi_2 - T\Phi_2|p_z| - |p_x|)}{2f_dR_oT\Phi_2}\frac{(L_y - |p_y|)}{L_y}. \tag{C.65}$$

As before, we optimize the viewpoint by setting $L_y = f_dR_o\Phi_x$ which gives us

$$TF_{rel} = \frac{(2f_dR_oT\Phi_2 - T\Phi_2|p_z| - |p_x|)}{2f_dR_oT\Phi_2}\frac{(f_dR_o\Phi_x - |p_y|)}{f_dR_o\Phi_x}. \tag{C.66}$$

We now wish to compute the statistics of TF_{rel} which is a function of the random variables p_x, p_y, p_z. First, we need the statistics of $|p_x|, |p_y|, |p_z|$. It is readily shown that

$$E[|p_x|] = \sigma_p\sqrt{\frac{2}{3\pi}}, \tag{C.67}$$

$$E[|p_x|^2] = \frac{{\sigma_p}^2}{3}, \tag{C.68}$$

$$E[|p_x||p_y|] = \frac{2{\sigma_p}^2}{3\pi}. \tag{C.69}$$

Given our model of pose error, the statistics of p_x, p_y and p_z are identical. Therefore, Equation C.66 can be written as

$$TF_{rel} = 1 - a|p_x| + b|p_x||p_y| \tag{C.70}$$

where

$$a = \frac{(\Phi_xT\Phi_2 + \Phi_x + 2T\Phi_2)}{2f_dR_o\Phi_xT\Phi_2} \tag{C.71}$$

and

$$b = \frac{(T\Phi_2 + 1)}{2{f_d}^2{R_o}^2\Phi_x T\Phi_2}. \tag{C.72}$$

Thus,

$$\begin{aligned} E[TF_{rel}] &= \mu_{TF_{rel}} \\ &= 1 - a\sigma_p\sqrt{\frac{2}{3\pi}} + \frac{2b{\sigma_p}^2}{3\pi}. \end{aligned} \tag{C.73}$$

Considering only terms up to quadratic

$${\mu_{TF_{rel}}}^2 \approx 1 + \frac{2a^2{\sigma_p}^2}{3\pi} - 2a\sigma_p\sqrt{\frac{2}{3\pi}} + \frac{4b{\sigma_p}^2}{3\pi} \tag{C.74}$$

and from Equation C.70

$${TF_{rel}}^2 \approx 1 + a^2|p_x|^2 - 2a|p_x| + 2b|p_x||p_y| \tag{C.75}$$

$$E[{TF_{rel}}^2] = 1 + \frac{a^2{\sigma_p}^2}{3} - 2a\sigma_p\sqrt{\frac{2}{3\pi}} + \frac{4b{\sigma_p}^2}{3\pi}. \tag{C.76}$$

Therefore,

$$\begin{aligned} {\sigma_{TF_{rel}}}^2 &= E[{TF_{rel}}^2] - {\mu_{TF_{rel}}}^2 \\ &= \frac{a^2{\sigma_p}^2(\pi - 2)}{3\pi}. \end{aligned} \tag{C.77}$$

In summary, the relative impact of pose position error on field of view erosion is approximately

$$\mu_{TF_{rel}} = 1 - \frac{(\Phi_x T\Phi_2 + \Phi_x + 2T\Phi_2)}{2f_d\Phi_x T\Phi_2}\sqrt{\frac{2}{3\pi}}\frac{\sigma_p}{R_o} + \frac{(T\Phi_2+1)}{3\pi {f_d}^2\Phi_x T\Phi_2}\frac{{\sigma_p}^2}{{R_o}^2} \tag{C.78}$$

$$\sigma_{TF_{rel}} = \frac{(\Phi_x T\Phi_2 + \Phi_x + 2T\Phi_2)}{2f_d\Phi_x T\Phi_2}\sqrt{\frac{\pi-2}{3\pi}}\frac{\sigma_p}{R_o}. \tag{C.79}$$

The impact of pose position error on viewpoint coverage calculated from Equations C.78 and C.79 is shown at Figure 7.9 for a line-scan sensor with $\Phi_x = 30°$, $f_d = 1.02$ and the optimum scanning geometry. Viewpoint coverage and effective field of view deteriorates for $TF_{rel} < 1$. By definition, TF_{rel} never exceeds one. One- and two-sigma curves for relative targeted footprint are displayed bracketing the average value. We observe that there is moderate frustum erosion with pose position error. Hence, it is an important factor in view planning.

Depth of Field Erosion As previously noted, pose error will result in erosion whenever $\delta_z > (f_d - 1)R_o$. With pose position error, such an event is catastrophic as the entire scan is compromised at the targeted stand-off distance. Therefore, the only statistic of interest relative to depth of field erosion is the probability $P[p_z > (f_d - 1)R_o]$. Then,

$$P[p_z > \epsilon] = 1 - \Phi[\frac{\epsilon - \mu}{\sigma_z}] \tag{C.80}$$

where $\sigma_z = \sigma_p/\sqrt{3}$, $\mu = 0$, $\epsilon = (f_d - 1)R_o$ and

$$\Phi[z] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z} e^{-u^2/2}\,du. \tag{C.81}$$

Then

$$P[p_z > (f_d - 1)R_o] = 1 - \Phi[\frac{\sqrt{3}(f_d - 1)}{(\sigma_p/R_o)}]. \tag{C.82}$$

The probability of depth of field erosion with pose position error is shown at Figure 7.7, plotted on a normalized logarithmic scale. These curves can be used to select a suitable distance factor f_d for a given level of pose error and tolerance for depth of field erosion.

C.2 Axis Error

C.2.1 Measurement Precision Impact

Precision Estimation Model This section considers the effect of the axis component of pose orientation error on measurement precision estimation. For the line-scan range camera examined here, we follow the convention that the frustum origin (which coincides with the laser rest position) falls half way along the y-axis linear scan (Figure 3.7). The sensor boresight is defined by the sensor negative z-axis. Axis error is modeled by a unit vector uniformly distributed on the surface of a cone centered on the camera boresight. Effectively, axis error involves rotations about the sensor x and y axes. These can be combined into a single rotation of angle α about an axis $\vec{r}$ in the xy-plane at angle β. The axis error cone half-angle α is modeled as a zero-mean Gaussian process with standard deviation σ_a while β, the angle in the xy-plane at which the axis error occurs, has a uniform distribution over $[-\pi, \pi]$.

The corresponding rotation matrix $R_{r,\alpha}$ is as follows [FGL87], where $V\alpha$ is $vers\,\alpha = 1 - \cos\alpha$:

$$\begin{pmatrix} {r_x}^2 V\alpha + C\alpha & r_x r_y V\alpha - r_z S\alpha & r_x r_z V\alpha + r_y S\alpha \\ r_x r_y V\alpha + r_z S\alpha & {r_y}^2 V\alpha + C\alpha & r_y r_z V\alpha - r_x S\alpha \\ r_x r_z V\alpha - r_y S\alpha & r_y r_z V\alpha + r_x S\alpha & {r_z}^2 V\alpha + C\alpha \end{pmatrix} \tag{C.83}$$

As the rotation vector $\vec{r}$ lies in the xy-plane, $\vec{r} = (r_x, r_y, r_z) = (S\beta, C\beta, 0)$. We can therefore simplify Equation C.83 to

$$R_{r,\alpha} = \begin{pmatrix} S^2\beta V\alpha + C\alpha & S\beta C\beta V\alpha & C\beta S\alpha \\ S\beta C\beta V\alpha & C^2\beta V\alpha + C\alpha & -S\beta S\alpha \\ -C\beta S\alpha & S\beta S\alpha & C\alpha \end{pmatrix} \tag{C.84}$$

It is again convenient to deal with relative precision P_{rel} from Equation 7.4

$$P_{rel} = \frac{z'^2 \cos\theta_{yz}}{z^2 \cos\theta'_{yz}}. \tag{C.85}$$

We assume the sensor range has been optimally set at $z = f_d R_o$. From Equation C.84,

$$z' = -C\beta S\alpha\, x + S\beta S\alpha\, y + C\alpha\, z. \tag{C.86}$$

Inclination angle θ_{yz} is the angle between the projection of the surface normal $\vec{n}$ on the y-z plane in camera space in the absence of pose axis error - that is, $\vec{n_{yz}} = (0, n_y, n_z)$, and the negative of the camera boresight $\vec{u_z} = (0,0,1)$. Similarly, inclination angle θ'_{yz} is the angle between the surface normal in camera space in the y-z plane in the presence of pose axis error $\vec{n_{yz}}' = (0, n_y', n_z')$ and $\vec{u_{z'}} = (0,0,1)$.

Here,

$$\begin{aligned} n_y{}' &= S\beta C\beta V\alpha n_x + (C^2\beta V\alpha + C\alpha)n_y - S\beta S\alpha n_z, \\ n_z{}' &= C\beta S\beta n_x + S\beta S\alpha n_y + C\alpha n_z. \end{aligned} \tag{C.87}$$

Then,

$$C\theta_{yz} = \frac{\vec{n_{yz}} \cdot \vec{u_z}}{||\vec{n_{yz}}||\ ||\vec{u_z}||}, \tag{C.88}$$

$$C\theta'_{yz} = \frac{\vec{n_{yz}}' \cdot \vec{u_{z'}}}{||\vec{n_{yz}}'||\ ||\vec{u_{z'}}||}. \tag{C.89}$$

Let us consider the case where the sensor is optimally positioned such that the inclination angles are zero. Therefore, $\theta_{xz} = \theta_{yz} = 0$ which implies $\vec{n} = (0, 0, 1)$. Additionally, we are concerned only with points near the sensor boresight - that is, $x = y = 0$. Then, $\vec{n_{yz}}' = (0, -S\beta S\alpha, C\alpha)$, $z'^2/z^2 = C^2\alpha$ and

$$C\theta_{yz} = 1, \tag{C.90}$$

$$C\theta'_{yz} = \frac{C\alpha}{\sqrt{S^2\beta S^2\alpha + C^2\alpha}}. \tag{C.91}$$

Then, the expression for relative precision P_{rel} at Equation C.85 is

$$\begin{aligned} P_{rel} &= C^2\alpha \frac{\sqrt{S^2\beta S^2\alpha + C^2\alpha}}{C\alpha} \\ &= C^2\alpha\sqrt{S^2\beta T^2\alpha + 1}. \end{aligned} \tag{C.92}$$

As $T^2\alpha \ll 1$, Equation C.92 can be written as follows

$$
\begin{aligned}
P_{rel} &= C^2\alpha(1 + \frac{S^2\beta T^2\alpha}{2} - \frac{S^4\beta T^4\alpha}{8}) \\
&= C^2\alpha + \frac{S^2\beta S^2\alpha}{2} - \frac{S^4\beta T^2\alpha S^2\alpha}{8}.
\end{aligned}
\tag{C.93}
$$

As a precursor to computing the statistics of P_{rel}, we next compute the statistics of α and β.

Statistics of β Recalling that β has a uniform distribution over $[-\pi, \pi]$, it can be shown that

$$E[S\beta] = E[C\beta] = E[S\beta C\beta] = 0 \tag{C.94}$$

$$E[S^2\beta] = E[C^2\beta] = 1/2 \tag{C.95}$$

$$E[S\beta C^2\beta] = E[S^2\beta C\beta] = E[S^3\beta] = 0 \tag{C.96}$$

$$E[S^2\beta C^2\beta] = 1/8 \tag{C.97}$$

$$E[S^4\beta] = 3/8 \tag{C.98}$$

$$E[|S\beta|] = E[|C\beta|] = 2/\pi \tag{C.99}$$

$$E[|S\beta C\beta|] = 1/\pi \tag{C.100}$$

Statistics of α As α is a zero-mean Gaussian process with standard deviation σ_a and neglecting high order terms, it can be shown that

$$E[S\alpha] = 0 \tag{C.101}$$

$$E[C\alpha] = e^{(-\frac{{\sigma_a}^2}{2})} \approx 1 - \frac{{\sigma_a}^2}{2} \tag{C.102}$$

$$E[T^2\alpha] = E[S^2\alpha] \approx {\sigma_a}^2 \tag{C.103}$$

$$E[C^2\alpha] \approx 1 - {\sigma_a}^2 \tag{C.104}$$

$$E[C^4\alpha] \approx 1 - 2{\sigma_a}^2 \tag{C.105}$$

$$E[S\alpha\, T\alpha] \approx {\sigma_a}^2 \tag{C.106}$$

$$E[S\alpha\, T^2\alpha] = 0 \tag{C.107}$$

$$E[S\alpha\, T^3\alpha] = E[S^2\alpha\, T^2\alpha] = \mathcal{O}({\sigma_a}^4) \tag{C.108}$$

$$E[|T\alpha|] = \sigma_a\sqrt{\frac{2}{\pi}} \tag{C.109}$$

$$E[S\alpha C\alpha] = 0 \tag{C.110}$$

$$E[S^2\alpha C^2\alpha] = {\sigma_a}^2 \tag{C.111}$$

Statistics of P_{rel} With the above results, we can now compute the statistics of P_{rel}. Applying the expected value operator to Equation C.93 and retaining terms up to quadratic in σ_a, we get the following approximation

$$\mu_{P_{rel}} = 1 - \frac{3{\sigma_a}^2}{4}. \tag{C.112}$$

Similarly, to compute the variance

$$(\mu_{P_{rel}})^2 \approx 1 - \frac{3{\sigma_a}^2}{2}, \tag{C.113}$$

$$\mu_{{P_{rel}}^2} = E[C^4\alpha + S^2\beta S^2\alpha C^2\alpha] = 1 - \frac{3{\sigma_a}^2}{2}. \tag{C.114}$$

So, using terms up to quadratic in σ_a, the variance ${\sigma_{P_{rel}}}^2$ is

$${\sigma_{P_{rel}}}^2 = \mu_{{P_{rel}}^2} - (\mu_{P_{rel}})^2 = 0. \tag{C.115}$$

Precision Impact Interpretation In summary, for the optimal viewing geometry (zero inclination angle and measurements near the sensor boresight), $\mu_{P_{rel}}$ and ${\sigma_{P_{rel}}}^2$ are as follows:

$$\mu_{P_{rel}} = 1 - \frac{3{\sigma_a}^2}{4}, \tag{C.116}$$

$${\sigma_{P_{rel}}}^2 = 0. \tag{C.117}$$

Average precision is marginally improved while the variance is zero. The first effect results from the small but consistent range reduction with axis error, while the second is due to the symmetrical viewing geometry under these conditions.

C.2.2 Sampling Density Impact

Sampling Density Estimation Model This section considers the effect of the axis component of pose orientation error on sampling density estimation. From Equation 7.9, our model for relative sampling density D_{rel} is

$$D_{rel} = \frac{C^2\theta'_{xz}C^2\theta'_{yz}}{C^2\theta_{xz}C^2\theta_{yz}} \frac{{R_{xz}}^2 C^2\theta_{yz} + {f_d}^2 {R_o}^2 C^2\theta_{xz}}{{R'_{xz}}^2 C^2\theta'_{yz} + {f_d}^2 {R_o}^2 C^2\theta'_{xz}}. \tag{C.118}$$

For axis error angle α at angle β to the xy-plane co-ordinate frame, we have the following geometric relationships:

$${R_{xz}}^2 = z^2 + x^2 = {f_d}^2 {R_o}^2 + x^2, \tag{C.119}$$

$${R'_{xz}}^2 = z'^2 + x'^2, \tag{C.120}$$

$$z' = -C\beta S\alpha \; x + S\beta S\alpha \; y + C\alpha \; z, \tag{C.121}$$

$$x' = (S^2\beta V\alpha + C\alpha) \; x + S\beta C\alpha V\alpha \; y + C\beta S\alpha \; z. \tag{C.122}$$

With the sensor in the optimal scanning configuration, $\theta_{xz} = \theta_{yz} = 0$ and $x = y = 0$. Then, as shown in the previous section

$$x' = C\beta S\alpha f_d R_o, \tag{C.123}$$

$$z' = C\alpha f_d R_o, \tag{C.124}$$

$${R_{xz}}^2 = {f_d}^2 {R_o}^2, \tag{C.125}$$

$${R'_{xz}}^2 = (C^2\beta S^2\alpha + C^2\alpha) {f_d}^2 {R_o}^2, \tag{C.126}$$

$$C\theta_{xz} = C\theta_{yz} = 1, \tag{C.127}$$

$$C\theta'_{xz} = \frac{C\alpha}{\sqrt{C^2\beta S^2\alpha + C^2\alpha}}, \tag{C.128}$$

$$C\theta'_{yz} = \frac{C\alpha}{\sqrt{S^2\beta S^2\alpha + C^2\alpha}}. \tag{C.129}$$

Using the above relationships, D_{rel} simplifies to

$$\begin{aligned} D_{rel} &= \frac{2C^2\alpha}{(C^2\beta S^2\alpha + C^2\alpha)^2 + (S^2\beta S^2\alpha + C^2\alpha)} \\ &= \frac{2}{1 + C^2\alpha + S^2\beta T^2\alpha + C^4\beta S^2\alpha T^2\alpha + 2C^2\beta S^2\alpha}. \end{aligned} \tag{C.130}$$

The expression for D_{rel} at Equation C.130 can be expressed in a Maclaurin series expansion in terms of α about $\alpha = 0$. Retaining only terms up to quadratic as usual, D_{rel} becomes

$$D_{rel} = 1 + \frac{(1 - 2C^2\beta - S^2\beta)\alpha^2}{2}. \tag{C.131}$$

Hence

$$\mu_{D_{rel}} = 1 - \frac{{\sigma_\alpha}^2}{4}. \tag{C.132}$$

Then

$$(\mu_{D_{rel}})^2 \approx 1 - \frac{{\sigma_\alpha}^2}{2}. \tag{C.133}$$

Similarly, a Maclaurin series expansion of ${D_{rel}}^2$ gives

$$ {D_{rel}}^2 = 1 + (1 - 2C^2\beta - S^2\beta)\alpha^2 \tag{C.134} $$

from which

$$ E[{D_{rel}}^2] = 1 - \frac{{\sigma_a}^2}{2} \tag{C.135} $$

so that

$$ {\sigma_{D_{rel}}}^2 = E[{D_{rel}}^2] - (\mu_{D_{rel}})^2 = 0. \tag{C.136} $$

Sampling Density Impact Interpretation In summary, for the optimal viewing geometry (zero inclination angle and measurements near the sensor boresight), $\mu_{D_{rel}}$ and ${\sigma_{D_{rel}}}^2$ are as follows:

$$ \mu_{D_{rel}} = 1 - \frac{{\sigma_a}^2}{4}, \tag{C.137} $$

$$ {\sigma_{D_{rel}}}^2 = 0. \tag{C.138} $$

Average sampling density is marginally reduced while the variance is zero. In the first case, benefits from the small but consistent range reduction are overpowered by the fact that all measurements under pose axis error are made slightly off boresight. As was the case for measurement precision, the variance of the sampling density is again zero due to the symmetrical viewing geometry under these conditions.

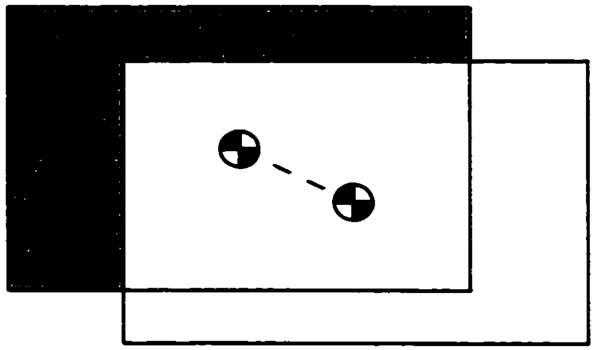


Figure C.3: Field of View Erosion with Axis Error

C.2.3 Visibility Impact

The visibility impact of sensor axis orientation error will depend on object shape, optical baseline length, sensor-object relative pose and the magnitude of the orientation error. Axis orientation error σ_a has the effect of causing a wobble in the position of the optical transmitter and receiver. The degree of wobble is determined by the orientation error, the length of the optical baseline and the position of the scan along the y-axis. Effects are similar to those generated by pose position error. Visibility impacts will generally be low at small axis orientation error σ_a.

C.2.4 Frustum Occupancy Impact

Field of View Erosion As illustrated in Figure C.3, axis orientation error about the sensor boresight changes the sensor aiming point and thus potentially has a large impact on actual versus planned coverage.

At the specified stand-off range of $f_d R_o$, the projection of the sensor boresight on the frustum cross-section is off-set by an amount $(d_x, d_y) =$

$(f_d R_o T\alpha C\beta, f_d R_o T\alpha S\beta)$. Recall that we model axis error with two components, α and β. The off-boresight angle error α is modeled as a zero-mean Gaussian process with standard deviation σ_a while the angular location around the boresight β has a uniform distribution over $[-\pi, \pi]$. Then, the targeted sensor footprint is reduced to the following, where the scan length has again been optimally configured:

$$TF' = (2f_d R_o T\Phi_2 - |d_x|)(f_d R_o \Phi_x - |d_y|). \tag{C.139}$$

Then, the relative targeted footprint is

$$\begin{aligned} TF_{rel} &= \frac{(2f_d R_o T\Phi_2 - f_d R_o|T\alpha C\beta|)}{2f_d R_o T\Phi_2} \frac{(f_d R_o \Phi_x - f_d R_o|T\alpha S\beta|)}{f_d R_o \Phi_x} \\ &= 1 - a|T\alpha C\beta| - b|T\alpha S\beta| + cT^2\alpha|C\beta S\beta| \end{aligned} \tag{C.140}$$

where $a = 1/2T\Phi_2$, $b = 1/\Phi_x$ and $c = 1/2T\Phi_2\Phi_x$. Consequently, the average relative targeted footprint is

$$\mu_{TF_{rel}} = 1 - \sigma_a(a+b)\frac{2}{\pi}\sqrt{\frac{2}{\pi}} + c\frac{{\sigma_a}^2}{\pi}. \tag{C.141}$$

To compute the variance, we have

$$(\mu_{TF_{rel}})^2 \approx 1 - \sigma_a(a+b)\frac{4}{\pi}\sqrt{\frac{2}{\pi}} + 2c\frac{{\sigma_a}^2}{\pi} + {\sigma_a}^2(a+b)^2\frac{8}{\pi^3}, \tag{C.142}$$

$$E[TF_{rel}{}^2] = 1 - \sigma_a(a+b)\frac{4}{\pi}\sqrt{\frac{2}{\pi}} + 2c\frac{{\sigma_a}^2}{\pi} + 2ab\frac{{\sigma_a}^2}{\pi} + \frac{{\sigma_a}^2(a^2+b^2)}{2}. \tag{C.143}$$

Thus

$$\sigma_{TF_{rel}}{}^2 = \frac{\sigma_a{}^2}{2\pi^3}[(a^2+b^2)(\pi^3-16)+4ab(\pi^2-8)]. \tag{C.144}$$

In summary, the relative impact of the axis component of pose orientation error on field of view erosion is approximately

$$\mu_{TF_{rel}} = 1 - \frac{(\Phi_x + 2T\Phi_2)}{\pi T\Phi_2\Phi_x}\sqrt{\frac{2}{\pi}}\sigma_a + \frac{\sigma_a{}^2}{2\pi T\Phi_2\Phi_x}, \tag{C.145}$$

$$\sigma_{TF_{rel}}{}^2 = \frac{\sigma_a{}^2}{8\pi^3T^2\Phi_2\Phi_x{}^2}[(4T^2\Phi_2+\Phi_x{}^2)(\pi^3-16)+8T\Phi_2\Phi_x(\pi^2-8)]. \tag{C.146}$$

The impact of the axis component of pose orientation error on viewpoint coverage calculated from Equations C.145 and C.146 is shown at Figure 7.10 for a line-scan sensor with $\Phi_x = 30°$ and the optimum scanning geometry. One- and two-sigma curves for relative targeted footprint are displayed bracketing the average value. We observe that there is a high level of frustum erosion with pose axis orientation error and, hence a serious impact on view planning.

Depth of Field Erosion Pose error produces frustum erosion whenever $\delta_z > (f_d - 1)R_o$. Unlike the case of pose position error which results in catastrophic frustum erosion of an entire scan, frustum erosion under axis error is a slowly occurring phenomena, starting at the far edges of the frustum cross-section and gradually moving inward with increasing error.

It is easily shown that depth of field erosion from axis error will begin to occur in the y-z plane when

$$S\alpha > \frac{f_d - 1}{f_d \Phi_2} \tag{C.147}$$

and in the x-z plane when

$$S\alpha > \frac{f_d - 1}{\Phi_2}. \tag{C.148}$$

For small α, $S\alpha \approx \alpha$. Thus, using the most conservative test from Equations C.147 and C.148, depth of field erosion will begin to occur for

$$\alpha > \frac{f_d - 1}{f_d \Phi_2}. \tag{C.149}$$

Therefore, the statistic of interest is the probability $P[\alpha > (f_d - 1)/(f_d \Phi_2)]$. Then,

$$P[\alpha > \tfrac{(f_d-1)}{f_d \Phi_2}] = 1 - \Phi[\tfrac{(f_d-1)}{f_d \Phi_2 \sigma_a}]. \tag{C.150}$$

The probability of depth of field erosion with pose axis error is shown at Figure 7.8, plotted on a normalized logarithmic scale. These curves can be compared with depth of field erosion with pose position error shown at Figure 7.7. Recall, however, that frustum depth of field erosion under axis error is gradual, whereas it is abrupt and catastrophic with position error. In summary, pose axis error results in only minor frustum depth of field erosion under most practical system configurations.

C.3 Twist Error

C.3.1 Measurement Precision Impact

Precision Estimation Model This section considers the effect of the twist component of pose orientation error on measurement precision estimation - that is, rotation error about the sensor boresight. Twist angle error α about the sensor z-axis is modeled as a zero-mean Gaussian process with standard deviation σ_t. The corresponding rotation matrix $R_{z,\alpha}$is

$$R_{z,\alpha} = \begin{pmatrix} C\alpha & -S\alpha & 0 \\ S\alpha & C\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{C.151}$$

Range remains constant $z' = z = f_d R_o$, so the relative measurement precision P_{rel} is

$$P_{rel} = \frac{\hat{\sigma}_{z'}}{\hat{\sigma}_z} = \frac{\cos\theta_{yz}}{\cos\theta'_{yz}}. \tag{C.152}$$

Inclination angle θ_{yz} is the angle between the projection of the surface normal $\vec{n}$ in camera space on the y-z plane $\vec{n_{yz}} = (0, n_y, n_z)$ and the camera boresight axis $\vec{u_z} = (0,0,1)$. Similarly, inclination angle θ'_{yz} is the angle between the perturbed surface normal $\vec{n_{yz}}' = (0, n_y', n_z') = (0, S\alpha\, n_x + C\alpha\, n_y, n_z)$ in camera space and $\vec{u_{z'}} = (0,0,1)$. Then,

$$C\theta_{yz} = \frac{\vec{n_{yz}} \cdot \vec{u_z}}{||\vec{n_{yz}}||\;||\vec{u_z}||} = \frac{n_z}{\sqrt{{n_y}^2 + {n_z}^2}}, \tag{C.153}$$

$$C\theta'_{yz} = \frac{\vec{n_{yz}}' \cdot \vec{u_{z'}}}{||\vec{n_{yz}}'||\;||\vec{u_{z'}}||} = \frac{n_z}{\sqrt{(S\alpha n_x + C\alpha n_y)^2 + {n_z}^2}}. \tag{C.154}$$

Consequently, as $n_x/n_z = T\theta_{xz}$ and $n_y/n_z = T\theta_{yz}$

$$\begin{aligned}
\frac{\cos\theta_{yz}}{\cos\theta'_{yz}} &= \sqrt{\frac{(S\alpha\, n_x + C\alpha\, n_y)^2 + {n_z}^2}{{n_y}^2 + {n_z}^2}} \\
&= \sqrt{\frac{(S\alpha T\theta_{xz} + C\alpha T\theta_{yz})^2 + 1}{T^2\theta_{yz} + 1}} \\
&= C\theta_{yz}\sqrt{1 + (S\alpha T\theta_{xz} + C\alpha T\theta_{yz})^2}. \qquad \text{(C.155)}
\end{aligned}$$

Then, the relative measurement precision P_{rel} is

$$P_{rel} = C\theta_{yz}\sqrt{1 + (S\alpha T\theta_{xz} + C\alpha T\theta_{yz})^2}. \qquad \text{(C.156)}$$

Statistics of P_{rel} Equation C.156 can be rewritten as follows, where $a = (S\alpha T\theta_{xz} + C\alpha T\theta_{yz})^2$:

$$P_{rel} = C\theta_{yz}\sqrt{1 + a}. \qquad \text{(C.157)}$$

As $T\theta_{xz} < \sqrt{3}$ and $T\theta_{yz} < \sqrt{3}$ in the range of interest (i.e. below threshold) and $\alpha \ll 1$, Equation C.157 can be approximated as

$$P_{rel} = C\theta_{yz}[1 + \frac{a}{2} - \frac{a^2}{8} + \frac{a^3}{16} - \frac{5a^4}{128} + \mathcal{O}(a^5)]. \qquad \text{(C.158)}$$

Therefore, using Equations [C.101-C.111] and taking terms up to quadratic in σ_t, we get the following:

$$\begin{aligned}
E[a] &= E[S^2\alpha T^2\theta_{xz} + C^2\alpha T^2\theta_{yz} + 2S\alpha C\alpha T\theta_{xz}T\theta_{yz}] \\
&= {\sigma_t}^2 T^2\theta_{xz} + (1 - {\sigma_t}^2)T^2\theta_{yz}, \qquad \text{(C.159)}
\end{aligned}$$

$$E[a^2] = 6\sigma_t{}^2T^2\theta_{xz}T^2\theta_{yz} + (1 - 2\sigma_t{}^2)T^4\theta_{yz}, \tag{C.160}$$

$$E[a^3] = 15\sigma_t{}^2T^2\theta_{xz}T^4\theta_{yz} + (1 - 3\sigma_t{}^2)T^6\theta_{yz}, \tag{C.161}$$

$$E[a^4] = 28\sigma_t{}^2T^2\theta_{xz}T^6\theta_{yz} + (1 - 4\sigma_t{}^2)T^8\theta_{yz}. \tag{C.162}$$

Collecting terms, we get

$$\mu_{P_{rel}} = C\theta_{yz}[s_1 - \frac{\sigma_t{}^2T^2\theta_{yz}}{2}s_2 + \frac{\sigma_t{}^2T^2\theta_{xz}}{2}s_3] \tag{C.163}$$

where

$$s_1 = 1 + \frac{1}{2}T^2 - \frac{1}{8}T^4 + \frac{1}{16}T^6 - \frac{5}{128}T^8 + \cdots, \tag{C.164}$$

$$s_2 = 1 - \frac{1}{2}T^2 + \frac{3}{8}T^4 - \frac{5}{16}T^6 + \frac{35}{128}T^8 + \cdots, \tag{C.165}$$

$$s_3 = 1 - \frac{3}{2}T^2 + \frac{15}{8}T^4 - \frac{35}{16}T^6 + \frac{315}{128}T^8 + \cdots \tag{C.166}$$

and $T = T\theta_{yz}$. Closed form solutions exist for these infinite series, as follows:

$$s_1 = [1 + T^2]^{1/2} = \frac{1}{C\theta_{yz}}, \tag{C.167}$$

$$s_2 = [1 + T^2]^{-1/2} = C\theta_{yz}, \tag{C.168}$$

$$s_3 = [1 + T^2]^{-3/2} = C^3\theta_{yz}. \tag{C.169}$$

So the average relative precision is

$$\mu_{P_{rel}} = 1 + \frac{\sigma_t{}^2}{2}[T^2\theta_{xz}C^4\theta_{yz} - S^2\theta_{yz}]. \tag{C.170}$$

To calculate the variance, we compute

$$\begin{aligned} E[P_{rel}{}^2] &= E[C^2\theta_{yz}(1+a)] \\ &= 1 + \sigma_t{}^2T^2\theta_{xz}C^2\theta_{yz} - \sigma_t{}^2S^2\theta_{yz} \end{aligned} \tag{C.171}$$

$$\mu_{P_{rel}}{}^2 \approx 1 + \sigma_t{}^2T^2\theta_{xz}C^4\theta_{yz} - \sigma_t{}^2S^2\theta_{yz} \tag{C.172}$$

from which the variance of P_{rel} is

$$\sigma_{P_{rel}}{}^2 = \sigma_t{}^2T^2\theta_{xz}S^2\theta_{yz}C^2\theta_{yz} \tag{C.173}$$

and the standard deviation is

$$\sigma_{P_{rel}} = \sigma_t T\theta_{xz}S\theta_{yz}C\theta_{yz}. \tag{C.174}$$

Precision Impact Interpretation Considering the behavior of relative precision P_{rel} for optimal scanning geometry of $\theta_{xz} = \theta_{yz} = 0$, we note from Equations C.170 and C.173 that its average is unchanged and variance is zero. These statistics result from the symmetrical viewing geometry under these special circumstances:

$$P_{rel} \approx 1, \tag{C.175}$$

$$\sigma_{P_{rel}} \approx 0. \tag{C.176}$$

Even with non-optimal scanning geometry, it is evident that the effect of twist orientation error on precision estimation is marginal.

C.3.2 Sampling Density Impact

Sampling Density Estimation Model This section considers the effect of the twist component of pose orientation error on sampling density estimation. From Equation 7.9, our model for relative sampling density D_{rel} is

$$D_{rel} = \frac{C^2\theta'_{xz}C^2\theta'_{yz}}{C^2\theta_{xz}C^2\theta_{yz}} \frac{R_{xz}{}^2 C^2\theta_{yz} + f_d{}^2 R_o{}^2 C^2\theta_{xz}}{R'_{xz}{}^2 C^2\theta'_{yz} + f_d{}^2 R_o{}^2 C^2\theta'_{xz}}. \tag{C.177}$$

For rotation error angle α about the boresight, we have the following geometric relationships

$$R_{xz}{}^2 = z^2 + x^2 = f_d{}^2 R_o{}^2 + x^2, \tag{C.178}$$

$$R'_{xz}{}^2 = z'^2 + x'^2, \tag{C.179}$$

$$z' = z = f_d R_o, \tag{C.180}$$

$$x' = C\alpha\, x - S\alpha\, y, \tag{C.181}$$

$$C^2\theta_{xz} = \frac{n_z{}^2}{n_x{}^2 + n_z{}^2}, \tag{C.182}$$

$$C^2\theta_{yz} = \frac{n_z{}^2}{n_y{}^2 + n_z{}^2}, \tag{C.183}$$

$$C^2\theta'_{xz} = \frac{n_z{}^2}{(C\alpha\, n_x - S\alpha\, n_y)^2 + n_z{}^2}, \tag{C.184}$$

$$C^2\theta'_{yz} = \frac{{n_z}^2}{(S\alpha\, n_x + C\alpha\, n_y)^2 + {n_z}^2}. \tag{C.185}$$

From the above, Equation C.177 reduces to

$$D_{rel} = \frac{(1 + {x_r}^2)(1 + T^2\theta_{xz}) + (1 + T^2\theta_{yz})}{(1 + {t_1}^2)(1 + {t_2}^2) + (1 + {t_3}^2)} = \frac{N}{D} \tag{C.186}$$

where $x_r = x/f_d R_o = \tan\phi_{xz}$, $y_r = y/f_d R_o$, $t_1 = C\alpha\, x_r - S\alpha\, y_r$, $t_2 = C\alpha T\theta_{xz} - S\alpha T\theta_{yz}$, and $t_3 = S\alpha T\theta_{xz} + C\alpha T\theta_{yz}$.

The numerator of Equation C.186 reduces to

$$\begin{aligned} N &= \frac{(1 + {x_r}^2)C^2\theta_{yz} + C^2\theta_{xz}}{C^2\theta_{xz}C^2\theta_{yz}} \\ &= \frac{C^2\theta_{yz} + C^2\phi_{xz}C^2\theta_{yz}}{C^2\theta_{yz}C^2\phi_{xz}C^2\theta_{yz}}. \end{aligned} \tag{C.187}$$

As $|t_1|$, $|t_2|$ and $|t_3|$ are all < 1 over the range of interest, the denominator can be written as $D = 2(1 + t)$ where $t = ({t_1}^2 + {t_2}^2 + {t_3}^2 + {t_1}^2{t_2}^2)/2$. Then, the relative sampling density D_{rel} is

$$D_{rel} = \frac{N}{2(1 + t)} \tag{C.188}$$

which can be further written as

$$D_{rel} = \frac{N}{2}(1 - t + \mathcal{O}(t^2)). \tag{C.189}$$

Statistics of D_{rel} Considering only terms up to quadratic in the key variables, the average relative sampling density $\mu_{D_{rel}}$ is approximately

$$\mu_{D_{rel}} = E[\tfrac{N}{2}(1-t)] \tag{C.190}$$

where

$$E[t] = \frac{1}{2}\{E[t_1{}^2] + E[t_2{}^2] + E[t_3{}^2] + E[t_1{}^2]E[t_2{}^2]\}, \tag{C.191}$$

$$\begin{aligned} E[t_1{}^2] &= E[C^2\alpha\, x_r{}^2 + S^2\alpha\, y_r{}^2 - 2S\alpha C\alpha\, x_r y_r] \\ &= (1-\sigma_t{}^2)x_r{}^2 + \sigma_t{}^2 y_r{}^2, \end{aligned} \tag{C.192}$$

$$\begin{aligned} E[t_2{}^2] &= E[C^2\alpha T^2\theta_{xz} + S^2\alpha T^2\theta_{yz} - 2S\alpha C\alpha T\theta_{xz}T\theta_{yz}] \\ &= (1-\sigma_t{}^2)T^2\theta_{xz} + \sigma_t{}^2 T^2\theta_{yz}, \end{aligned} \tag{C.193}$$

$$\begin{aligned} E[t_3{}^2] &= E[S^2\alpha T^2\theta_{xz} + C^2\alpha T^2\theta_{yz} + 2S\alpha C\alpha T\theta_{xz}T\theta_{yz}] \\ &= \sigma_t{}^2 T^2\theta_{xz} + (1-\sigma_t{}^2)T^2\theta_{yz}, \end{aligned} \tag{C.194}$$

$$\begin{aligned} E[t_1{}^2 t_2{}^2] &= E[(C^2\alpha\, x_r{}^2 + S^2\alpha\, y_r{}^2 - 2S\alpha C\alpha\, x_r y_r) \\ &\quad (C^2\alpha T^2\theta_{xz} + S^2\alpha T^2\theta_{yz} - 2S\alpha C\alpha T\theta_{xz}T\theta_{yz})] \\ &= (1-2\sigma_t{}^2)x_r{}^2 T^2\theta_{xz} + \sigma_t{}^2 y_r{}^2 T^2\theta_{xz} \\ &\quad + \sigma_t{}^2 x_r{}^2 T^2\theta_{yz} + 4\sigma_t{}^2 x_r y_r T\theta_{xz}T\theta_{yz}, \end{aligned} \tag{C.195}$$

$$
\begin{aligned}
E[t] &= \frac{1}{2}\{{x_r}^2(1+T^2\theta_{xz}) + T^2\theta_{xz} + T^2\theta_{yz} + \\
&\quad {\sigma_t}^2[-{x_r}^2(1+2T^2\theta_{xz}) + {y_r}^2(1+T^2\theta_{xz}) + {x_r}^2T^2\theta_{yz} + 4x_ry_rT\theta_{xz}T\theta_{yz}]\}.
\end{aligned}
\tag{C.196}
$$

Collecting terms, we get the following expression for the average relative sampling density $\mu_{D_{rel}}$, where N is defined at Equation C.187:

$$
\begin{aligned}
\mu_{D_{rel}} &\approx \frac{N}{4}[2 - {x_r}^2(1+T^2\theta_{xz}) - T^2\theta_{xz} - T^2\theta_{yz}] \\
&\quad + \frac{{\sigma_t}^2N}{4}[{x_r}^2(1+2T^2\theta_{xz}) - {y_r}^2(1+T^2\theta_{xz}) - {x_r}^2T^2\theta_{yz} - 4x_ry_rT\theta_{xz}T\theta_{yz}].
\end{aligned}
\tag{C.197}
$$

For the variance, we have the following, considering terms to $\mathcal{O}(t^2)$

$$\mu_{D_{rel}} \approx \frac{N}{2}(1 - E[t] + E[t^2]), \tag{C.198}$$

$${\mu_{D_{rel}}}^2 \approx \frac{N^2}{4}(1 - 2E[t] + (E[t])^2 + 2E[t^2]), \tag{C.199}$$

$$E[{D_{rel}}^2] = \frac{N^2}{4}(1 - 2E[t] + 3E[t^2]). \tag{C.200}$$

Therefore, the variance ${\sigma_{D_{rel}}}^2$ is

$$
\begin{aligned}
{\sigma_{D_{rel}}}^2 &\approx \frac{N^2}{4}(E[t^2] - (E[t])^2) \\
&= \frac{N^2{\sigma_t}^2{x_r}^2}{4}[y_r(1+T^2\theta_{xz}) + x_rT\theta_{xz}T\theta_{yz})]^2.
\end{aligned}
\tag{C.201}
$$

Hence, the standard deviation $\sigma_{D_{rel}}$ is

$$\sigma_{D_{rel}} \approx \frac{N\sigma_t x_r}{2}[y_r(1 + T^2\theta_{xz}) + x_r T\theta_{xz} T\theta_{yz})]. \tag{C.202}$$

Sampling Density Impact Interpretation Considering the behavior of relative sampling density for optimal scanning geometry of $\theta_{xz} = \theta_{yz} = 0$, the coefficient N reduces to $(1 + C^2\phi_{xz})/C^2\phi_{xz}$ and we obtain the following simplified expressions for the average and variance of D_{rel} from Equations C.197 and C.202:

$$\mu_{D_{rel}} \approx \frac{N}{4}[2 - x_r{}^2] + \frac{\sigma_t{}^2 N}{4}[x_r{}^2 - y_r{}^2], \tag{C.203}$$

$$\sigma_{D_{rel}} \approx \frac{N\sigma_t x_r y_r}{2}. \tag{C.204}$$

Furthermore, for measurements near the sensor boresight $(x_r \approx 0 \approx y_r)$, then

$$\mu_{D_{rel}} \approx 1, \tag{C.205}$$

$$\sigma_{D_{rel}} \approx 0. \tag{C.206}$$

Even with non-optimal scanning geometry, it is evident that the effect of twist orientation error on sampling density estimation is marginal.

C.3.3 Visibility Impact

The visibility impact of sensor twist orientation error will depend on the object shape, sensor optical baseline, sensor-object relative pose and the

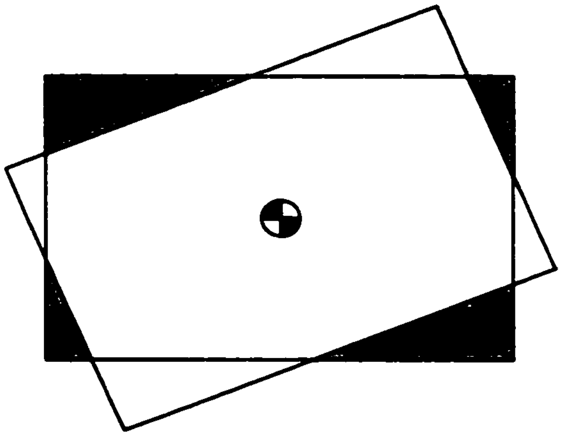


Figure C.4: Field of View Erosion with Twist Error

magnitude of the orientation error. The impact on visibility will generally be low at small twist orientation error σ_t. However, occasionally even small orientation errors may be sufficient to push the visibility of surface elements near threshold into occlusion. Moderate to large twist orientation errors will result in significant unplanned occlusions due to shadow effects arising from the non-zero optical baseline.

C.3.4 Frustum Occupancy Impact

Field of View Erosion As illustrated in Figure C.4, twist orientation error about the sensor boresight reduces the effective coverage. It is readily shown that coverage reduction δ_{TF} is as follows, where W and L are the width and length of the frustum cross-section at the desired stand-off range:

$$\delta_{TF} = \frac{1 - C\alpha}{2|S\alpha|C\alpha}[L^2 + W^2 - 2LW|S\alpha|]. \tag{C.207}$$

As we are interested only in relatively small angle errors, the absolute value operator need only be applied to the $S\alpha$ term. It is convenient to rewrite the foregoing as

$$\delta_{TF} = \frac{LW(C\alpha - 1)}{C\alpha} + \frac{|S\alpha|(L^2 + W^2)}{2C\alpha(1 + C\alpha)}. \tag{C.208}$$

For small α and considering only terms up to quadratic in the series expansion for $C\alpha$, Equation C.208 can be approximated by the following:

$$\delta_{TF} \approx -\frac{LW\alpha^2}{2} + \frac{|\alpha|(L^2 + W^2)}{4}. \tag{C.209}$$

Consequently, the relative targeted footprint is

$$\begin{aligned} TF_{rel} &= \frac{LW - \delta_{TF}}{LW} \\ &= 1 - \frac{|\alpha|(L^2 + W^2)}{4LW} + \frac{\alpha^2}{2}. \end{aligned} \tag{C.210}$$

As the frustum width at the optimum stand-off range is $W = 2f_d R_o T\Phi_2$ and the scan length has been optimized to $L = f_d R_o \Phi_x$, Equation C.210 is

$$TF_{rel} = 1 - a|\alpha| + \frac{\alpha^2}{2} \tag{C.211}$$

where

$$a = \frac{{\Phi_x}^2 + 4T^2\Phi_2}{8\Phi_x T\Phi_2}. \tag{C.212}$$

Then, the average relative targeted footprint is

$$\mu_{TF_{rel}} = 1 - a\sigma_t\sqrt{\frac{2}{\pi}} + \frac{{\sigma_t}^2}{2}. \tag{C.213}$$

Similarly,

$$E[TF_{rel}{}^2] = 1 - 2a\sigma_t\sqrt{\frac{2}{\pi}} + (a^2+1){\sigma_t}^2. \tag{C.214}$$

Consequently, the variance $\sigma_{TF_{rel}}$ is

$$\begin{aligned} {\sigma_{TF_{rel}}}^2 &= E[TF_{rel}{}^2] - (\mu_{TF_{rel}})^2 \\ &= a^2{\sigma_t}^2\frac{(\pi-2)}{\pi}. \end{aligned} \tag{C.215}$$

In summary, the relative impact of the twist component of viewpoint orientation error on field of view erosion is approximately

$$\mu_{TF_{rel}} = 1 - \frac{{\Phi_x}^2 + 4T^2\Phi_2}{8\Phi_x T\Phi_2}\sqrt{\frac{2}{\pi}}\sigma_t + \frac{{\sigma_t}^2}{2}, \tag{C.216}$$

$$\sigma_{TF_{rel}} = \frac{{\Phi_x}^2 + 4T^2\Phi_2}{8\Phi_x T\Phi_2}\sqrt{\frac{\pi-2}{\pi}}\sigma_t. \tag{C.217}$$

The impact of the twist component of pose orientation error on viewpoint coverage calculated from Equations C.216 and C.217 is shown in Figure 7.11 for a line-scan sensor with $\Phi_x = 30°$. One- and two-sigma curves for relative targeted footprint are displayed bracketing the average value. We observe that pose twist orientation error results in low frustum erosion and thus impacts view planning.

Depth of Field Erosion As twist orientation error does not change the sensor stand-off range, it does not erode the sensor depth of field.

Bibliography

[Abr97] Steven Abrams. *Sensor Planning in an Active Robot Work-Cell.* PhD thesis, Columbia University, 1997.

[AF99] Tal Arbel and Frank Ferrie. Entropy-based gaze planning. In *Proc. 2nd IEEE Workshop on Perception for Mobile Agents, in association with 1999 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, Fort Collins, Colorado*, pages 87–94, June 1999.

[Arb00] Tal Arbel. *Active Object Recognition Conditioned by Probabilistic Evidence and Entropy Maps.* PhD thesis, McGill University, 2000.

[Aub94] Stephane Aubry. *3D Model Construction from Multiple Sensor Viewpoints.* PhD thesis, McGill University, 1994.

[BA96] J.E. Banta and M.A. Abidi. Autonomous placement of a range sensor for acquisition of optimal 3D models. In *Proc. 1996 IEEE IECON. 22nd Int. Conf. on Industrial Electronics, Control and Instrumentation, Taipei, Taiwan*, pages 1583–1588, August 5-10 1996.

[Ban96] Joseph E. Banta. Optimal range sensor positioning for three-dimensional model reconstruction. Master's thesis, Univ. of Tennessee, Knoxville, 1996.

[Bat89] B.G. Batchelor. A prolog lighting advisor. In *Proc. SPIE Intell. Robots and Computer Vision VIII: Systems and Applications*, volume 1193, pages 295–302, 1989.

[BC96] J.E. Beasley and P.C. Chu. A genetic algorithm for the set covering problem. *European Journal of Operational Research*, 94:392–404, 1996.

[BD90] K.W. Bowyer and C.R. Dyer. Aspect graphs : An introduction and survey of recent results. *SPIE Close-Range Photogrammetry Meets Machine Vision*, 1395:200–208, 1990.

[Bea90] J.E. Beasley. A lagrangian heuristic for set covering problems. *Naval Research Logistics*, 37:151–164, 1990.

[BEC93] J.-A. Beraldin, S.F. El-Hakim, and L. Cournoyer. Practical range camera calibration. In *Proc. Videometrics II, SPIE, Boston, Massachusetts*, volume 2067, pages 21–30, 9-10 September 1993.

[Bes89] Paul J. Besl. Range image sensors. In J. Sanz, editor, *Advances in Machine Vision*. Springer-Verlag, New York, 1989.

[Bes96] Leland C. Best. *Autonomous Construction of Three-Dimensional Models from Range Data*. PhD thesis, Univ. of Wyoming, 1996.

[BM92] Paul J. Besl and Neil D. McKay. A method for registration of 3D shapes. *IEEE Trans. PAMI*, 14(2):239–256, February 1992.

[Buz90] Michael Joseph Buzinski. A comparison of touch probe and laser triangulation sensors for verifying the dimensional accuracy of manufactured parts. Master's thesis, Purdue University, 1990.

[BYW+95] J.E. Banta, Y.Zhien, X.Z. Wang, G. Zhang, M.T. Smith, and M.A. Abidi. A best-next-view algorithm for three-dimensional scene reconstruction using range images. *SPIE*, 2588:418–429, 1995.

[CB89] C.K. Cowan and A. Bergman. Determining the camera and light source location for a visual task. In *Proc. IEEE Int. Conf. on Robotics and Automation*, pages 509–514, 1989.

[CK88] C.K. Cowan and P.D. Kovesi. Automatic sensor placement from vision task requirements. *IEEE Trans. PAMI*, 10(3):407–416, May 1988.

[CL96] Brian Curless and Marc Levoy. Better optical triangulation through space-time analysis. In *SIGGRAPH '96*, pages 1–10, 1996.

[CMKD96] P. Cauchick-Miguel, T. King, and J. Davis. CMM verification: A survey. *Measurement*, 17(1):1–16, 1996.

[Con85] C.I. Connolly. The determination of next best views. In *Proc. IEEE Int. Conf. on Robotics and Automation*, pages 432–435, 1985.

[Cur97] Brian Lee Curless. *New Methods for Surface Reconstruction from Range Images*. PhD thesis, Stanford Univ., 1997.

[Dig] Adaptive scan. Technical report, Digibotics Ltd., www.digibotics.com.

[EB94] S.F. El-Hakim and J.-A. Beraldin. On the integration of range and intensity data to improve vision-based three-dimensional measurements. In *Proc. Videometrics III, SPIE, Boston, Massachusetts*, volume 2350, pages 306–321, 2-4 November 1994.

[EB95] S.F. El-Hakim and J.-A. Beraldin. Configuration design for sensor integration. In *Proc. Videometrics IV, SPIE, Philadelphia, Pennsylvania*, volume 2598, pages 274–285, 25-26 October 1995.

[FGL87] K.S. Fu, R.C. Gonzalez, and C.S.G. Lee. *Robotics: Control, Sensing, Vision and Intelligence*. McGraw-Hill, 1987.

[FMA+92] Olivier Faugeras, Joe Mundy, Narendra Ahuja, Charles Dyer, Alex Pentland, Ramesh Jain, and Katsushi Ikeuchi. Why aspect graphs are not (yet) practical for computer vision. *CVGIP: Image Understanding*, 55(2):212–218, 1992.

[FW82] M. Fisher and L. Wolsey. On the greedy heuristic for covering and packing problems. *SIAM Journal of Algebraic and Discrete Methods*, 3(4):584–591, Dec 1982.

[GJ79] M.R. Garey and D.S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W.H. Freeman, 1979.

[Gre99] Michael Greenspan. *Geometric Probing for 3D Object Recognition in Dense Range Data.* PhD thesis, Carleton University, 1999.

[GVS98] M.A. Garcia, S. Velazquez, and A.D. Sappa. A two-stage algorithm for planning the next view from range images. In *British Machine Vision Conference 1998*, pages 720–729, 1998.

[GVSB98] M.A. Garcia, S. Velazquez, A.D. Sappa, and L. Basanez. Autonomous sensor planning for 3D reconstruction of complex objects from range images. In *Proc. IEEE Int. Conf. on Robotics and Automation, Leuven, Belgium*, pages 3085–3090, 1998.

[GW97] T. Grossman and A. Wool. Computational experience with approximation algorithms for the set covering problem. *European Journal of Operational Research*, 101:81–92, 1997.

[HB94] Donald Hearn and M. Pauline Baker. *Computer Graphics.* Prentice-Hall Inc., 1994.

[HH98] Olaf Hall-Holt. www-graphics.stanford.edu/ olaf/nbv.html. Technical report, Stanford Univ., 1998.

[HK89] Seth A. Hutchinson and Avinash C. Kak. Planning sensing strategies in a robot work cell with multi-sensor capabilities. *IEEE Trans. Robotics and Automation*, 5(6):765–783, December 1989.

[Hym] Rapidscan. Technical report, Hymarc Ltd., www.hymarc.com.

[Jun99] Dieter Jungnickel. *Graphs, Networks and Algorithms.* Springer. 1999.

[KKK80] J. Kahn, M. Klawe, and D. Kleitman. Traditional galleries require fewer watchmen. Technical Report RJ3021. IBM Watson Research Center, NY, 1980.

[KS00] Kiriakos N. Kutulakos and Steven M. Seitz. A theory of shape by space carving. *Intl. Jour. Comp. Vis.*, 38(3):1999–218, 2000.

[KST90] Y. Kitamura, H. Sato, and H. Tamura. An expert system for industrial machine vision. In *Proc. 10th Int. Conf. on Pattern Recognition*, pages 771–773, 1990.

[Kut94] Kirialos N. Kutulakos. *Exploring Three-Dimensional Objects by Controlling the Point of Observation.* PhD thesis, Univ. of Wisconsin-Madison, 1994.

[LBG99] D.G. Lamb, D.L. Baird, and M.A. Greenspan. An automation system for industrial 3D laser digitizing. In *Second Int. Conf. on 3D Digital Imaging and Modeling, Ottawa*, pages 148–157, October 4-8 1999.

[LGG01] Christian Langis, Michael Greenspan, and Guy Godin. The parallel iterative closest point algorithm. In *Third Int. Conf. on 3D Digital Imaging and Modeling, Quebec City, Canada*, pages 195–202, May 2001.

[Mas97] Nikos Massios. Predicting the best next view for a laser range striper system. Master's thesis, Univ. of Edinburugh, 1997.

[Mav95] Jasna Maver. *Collecting Visual Information Using an Active Sensor System.* PhD thesis, University of Ljubljana, 1995.

[MB90] Jasna Maver and Ruzena Bajcsy. How to decide form the first view where to look next. In *Proc. DARPA Image Understanding Workshop, Pittsburgh, Pennslyvania*, pages 482–496, September 1990.

[MB93] Jasna Maver and Ruzena Bajcsy. Occlusions as a guide for planning the next view. *IEEE Trans. PAMI*, 17(5):417–433, May 1993.

[MBV96] M.J. Milroy, C. Bradley, and G.W. Vickers. Automated laser scanning based on orthogonal cross sections. *Machine Vision and Applications*, 9:106–118, 1996.

[MF98] Nickolaos A. Massios and Robert B. Fisher. A best next view selection algorithm incorporating a quality criterion. In *British Machine Vision Conference 1998*, pages 780–789, September, 1998.

[MTY01] Takashi Machida, Haruo Takemura, and Naokazu Yokoya. Dense estimation of surface reflectance parameters by selecting optimum illumination conditions. In *14th Int. Conf. on Vision Interface, Ottawa*, pages 244–251, 2001.

[MZH99] Ken'ichi Morooka, Hongbin Zha, and Tsutomu Hasegawa. Computations on a spherical view space for efficient planning of viewpoints in 3D object modeling. In *Second Int. Conf. on 3D Digital Imaging and Modeling, Ottawa*, pages 138–147, October 4-8 1999.

[New93] T.S. Newman. *Experiments in 3D CAD-based inspection using range images*. PhD thesis, Michigan State University, 1993.

[New00] Newport. Motion control catalogue. Technical report, www.newport.com, 2000.

[NJ95] T.S. Newman and A.K. Jain. A survey of automated visual inspection. *Computer Vision and Image Understanding*, 61(2):231–262, March 1995.

[Nov88] A. Novini. Lighting and optics expert system for machine vision. In *Proc. SPIE Optics, Illumination, and Image Sensing for Machine Vision III*, volume 1005, pages 131–136, 1988.

[PB95] Richard Pito and Ruzena Bajcsy. A solution to the next-best-view problem for automated cad model acquisition of free-form objects using range cameras. In *SPIE Symposium on Modeling, Simulation and Control Technologies for Manufacturing, Philadelphia, Penn.*, volume 2596, pages 78–89, October 25-26 1995.

[Pit96a] Richard Pito. Mesh integration based on co-measurements. In *Int. Conf. on Image Processing, Lausanne, Switzerland*, volume 2, pages 397–400, September 1996.

[Pit96b] Richard Pito. A sensor based solution to the next best view problem. In *Proc. Int. Conf. on Pattern Recognition*, pages 941–945, August 1996.

[Pit97a] Richard Pito. *Automated Surface Acquisition Using Range Cameras.* PhD thesis, Univ. of Pennsylvania, 1997.

[Pit97b] Richard Pito. A registration aid. In *Int. Conf. on Recent Advances in 3D Digital Imaging and Modeling, Ottawa*, pages 85–92, May 12-15 1997.

[Pit99] Richard Pito. A solution to the next best view problem for automated surface acquisition. *IEEE Trans. PAMI*, 21(10):1016–1030, October 1999.

[PO97] Dimitri Papadopoulos-Orfanos. *Numerisation geometrique automatique a l'aide d'une camera 3D de precision a profondeur de champ reduite.* PhD thesis, Ecole nationale superieure des telecommunications, Paris, France, 1997.

[POS97] Dimitri Papadopoulos-Orfanos and Francis Schmitt. Automatic 3D digitization using a laser rangefinder with a small field of view. In *Int. Conf. on Recent Advances in 3D Digital Imaging and Modeling, Ottawa*, pages 60–67, May 12-15 1997.

[PRBL99] F. Prieto, T. Redarce, P. Boulanger, and R. Lepage. CAD-based range sensor placement for optimum 3D data acquisition. In *Second Int. Conf. on 3D Digital Imaging and Modeling, Ottawa*, pages 128–137, 4-8 October 1999.

[PRBL01] F. Prieto, T. Redarce, P. Boulanger, and R. Lepage. Tolerance control with high resolution 3D measurements. In *Third Int. Conf. on 3D Digital Imaging and Modeling, Quebec City, Canada*, pages 339–346, 28 May - 1 June 2001.

[Pri99] Flavio Prieto. *Métrologie assistée par ordinateur: Apport des capteurs 3D sans contact.* PhD thesis, L'Institut National des Sciences Appliquées de Lyon, 1999.

[PRLB99] F. Prieto, T. Redarce, R. Lepage, and P. Boulanger. Range image accuracy improvement by acquisition planning. In *Int. Conf. on Vision Interface, Trois-Rivieres*, pages 156–163, 19-21 May 1999.

[Pud94] Christopher John Pudney. *Surface Modeling and Surface Following for Robots Equipped with Range Sensors.* PhD thesis, Univ. of Western Australia, Perth, 1994.

[RA92] Besma Roui-Abidi. *Automatic Sensor Placement for Volumetric Object Characterization.* PhD thesis, Univ. of Tennessee, Knoxville, 1992.

[RA97] Michael K. Reed and Peter K. Allen. A robotic system for 3D model acquisition from multiple range images. In *Proc. IEEE Int. Conf. on Robotics and Automation, Albuquerque, New Mexico,* pages 2509–2514, April 1997.

[RA99] M. Reed and P.K. Allen. 3D modeling from range imagery. *Image and Vision Computing,* 17(1):99–111, February 1999.

[RAS97a] M.K. Reed, P.K. Allen, and I. Stamos. 3D modeling from range imagery: An incremental method with a planning component. In *Int. Conf. on Recent Advances in 3D Digital Imaging and Modeling, Ottawa,* pages 76–83, May 12-15 1997.

[RAS97b] M.K. Reed, P.K. Allen, and I. Stamos. Automated model acquisition from range images with view planning. In *Proc. IEEE Conf. Vis. Pat. Rec., San Juan, Puerto Rico,* pages 72–77, June 17-19 1997.

[Ree93] Colin R. Reeves. *Modern Heuristic Techniques for Combinatorial Problems.* Blackwell Scientific Publications, Oxford, 1993.

[Ree98] Michael K. Reed. *Solid Model Acquisition form Range Imagery.* PhD thesis, Columbia University, 1998.

[RL01] Szymon Rusinkiewicz and Marc Levoy. Efficient variants of the ICP algorithm. In *Third Int. Conf. on 3D Digital Imaging and Modeling, Quebec City, Canada,* pages 145–152, May 2001.

[RM97] D.R. Roberts and A.D. Marshall. A review of viewpoint planning. Technical Report 97007, Univ. of Wales Cardiff, 1997.

[Rot00] Gerhard Roth. Building models from sensor data: an application shared by the computer vision and the computer graphics community. In A. Leonardis et al, editor, *Confluence of Computer Vision and Computer Graphics.* Kluwer Academic Publishers, Netherlands, 2000.

[RW97] G. Roth and W. Wibowo. An efficient volumetric method for building closed triangular meshes from 3D image and point data. In *Proc. Graphics Interface Conf. 97*, May 1997.

[SCF98] G. Soucy, F.G. Callari, and F.P. Ferrie. Uniform and complete surface coverage with a robot-mounted laser rangefinder. *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, Victoria, B.C., Canada*, pages 1682–1688, October 1998.

[Sco02] W.R. Scott. Pose error effects on range sensing. In *Submitted to 15th Int. Conf. on Vision Interface, Calgary*, 2002.

[Sen93] S. Sen. Minimal cost set covering using probabilistic methods. In *ACM Symp. Applied Computing, Indianapolis*, pages 157–164, 1993.

[Set96] J.A. Sethian. *Level Set Methods.* Cambridge University Press, 1996.

[SG93] Sergio W. Sedas-Gersey. *Algorithms for Automatic Sensor Placement to Acquire Complete and Accurate Information.* PhD thesis, Carnegie Mellon Univ., 1993.

[SNSS92] S. Sakane, R. Niepold, T. Sato, and Y. Shirai. Illumination setup planning for a hand-eye system based on an environmental model. *Advanced Robotics*, 6(4):461–482, 1992.

[SRR00] W.R. Scott, G. Roth, and J.-F. Rivest. Performance-oriented view planning for automatic model acquisition. In *31st Int. Sym. on Robotics, Montreal*, pages 314–319, May 2000.

[SRR01a] W.R. Scott, G. Roth, and J.-F. Rivest. View planning as a set covering problem. Technical Report NRC-44892, National Research Council of Canada, Institute for Information Technology, August 2001.

[SRR01b] W.R. Scott, G. Roth, and J.-F. Rivest. View planning for multi-stage object reconstruction. In *14th Int. Conf. on Vision Interface, Ottawa*, pages 64–71, June 2001.

[SRR01c] W.R. Scott, G. Roth, and J.-F. Rivest. View planning with a registration constraint. In *Third Int. Conf. on 3D Digital Imaging and Modeling, Quebec City, Canada*, pages 127–134, May 2001.

[SRR01d] W.R. Scott, G. Roth, and J.-F. Rivest. View planning with pose error. Technical Report NRC-44195, National Research Council of Canada, Institute for Information Technology, May 2001.

[SRR01e] W.R. Scott, G. Roth, and J.-F. Rivest. View planning with pose uncertainty. *Submitted to Computer Vision and Image Understanding*, 2001.

[STS92] J.A. Soons, F.C. Thomas, and P.H. Schellekens. Modeling the errors of multi-axis machines: A general methodology. *Precision Engineering*, 14(1):5–19, July 1992.

[SW92] Y.C. Shin and Y. Wei. A statistical analysis of positional errors of a multi-axis machine tool. *Precision Engineering*, 14(3):139–146, July 1992.

[Tar92] Konstantinos Tarabanis. *Sensor Planning and Modeling for Machine Vision Tasks*. PhD thesis, Columbia University, 1992.

[Tar93] Glen H. Tarbox. *A Volumetric Approach to Automated 3D Inspection*. PhD thesis, Rensselaer Polytechnic Institute, Ann Arbor, MI, 1993.

[TAT95] Konstantinos Tarabanis, Peter K. Allen, and Roger Y. Tsai. A survey of sensor planning in computer vision. *IEEE Trans. Robotics and Automation*, 11(1):86–104, February 1995.

[TG95] G.H. Tarbox and S.N. Gottschlich. Planning for complete sensor coverage in inspection. *Computer Vision and Image Understanding*, 61(1):84–111, January 1995.

[TTA95] Konstantinos Tarabanis, Roger Y. Tsai, and Peter K. Allen. The MVP sensor planning system for robotic vision tasks. *IEEE Trans. Robotics and Automation*, 11(1):72–85, February 1995.

[TTK96] Konstantinos Tarabanis, Roger Y. Tsai, and Anil Kaul. Computing occlusion-free viewpoints. *IEEE Trans. PAMI*, 18(3):279–292, March 1996.

[TUWR97] E. Trucco, M. Umasuthan, A.M. Wallace, and V. Roberto. Model-based planning of optimal sensor placements for inspection. *IEEE Trans. Robotics and Automation*, 13(2):182–194, April 1997.

[Urr00] Jorge Urrutia. Art gallery and illumination problems. In J.-R. Sack and J. Urrutia, editors, *Handbook of Computational Geometry*. Elsevier, New York, 2000.

[WF90] Peter Whaite and Frank P. Ferrie. From uncertainty to visual exploration. In *Third Int. Conf. on Computer Vision, Osaka, Japan*, pages 690–697, December 1990.

[WF91] Peter Whaite and Frank P. Ferrie. From uncertainty to visual exploration. *IEEE Trans. PAMI*, 13(10):1038–1049, October 1991.

[WF92] Peter Whaite and Frank P. Ferrie. Uncertain views. In *Proc. IEEE Conf. Vis. Pat. Rec.*, pages 3–9, June 15-18 1992.

[WF97] Peter Whaite and Frank P. Ferrie. Autonomous exploration: Driven by uncertainty. *IEEE Trans. PAMI*, 19(3):193–205, March 1997.

[Whe96] Mark Damon Wheeler. *Automatic Modeling and Localization for Object Recognition*. PhD thesis, Carnegie Mellon Univ., 1996.

[Win95] Wayne L. Winston. *Introduction to Mathematical Programming*. Duxbury Press, 1995.

[Wix94] Lambert Ernest Wixson. *Gaze Selection for Visual Search*. PhD thesis, Univ. of Rochester, 1994.

[XCB86] S. Xie, T. Calvert, and B. Bhattacharya. Planning views for the incremental construction of model bodies. In *Proc. 8th Int. Conf. on Pattern Recognition, Paris, France*, pages 154–157, October 27-31 1986.

[Ye97] Yiming Ye. *Sensor Planning for Object Search*. PhD thesis, Univ. of Toronto, 1997.

[YHS90] Seungku Yi, Robert M. Haralick, and Linda G. Shapiro. Automatic sensor and light source positioning for machine vision. In *Proc. 10th Int. Conf. on Pattern Recognition*, pages 55–59, 1990.

[YHS95] Seungku Yi, Robert M. Haralick, and Linda G. Shapiro. Optimal sensor and light source positioning for machine vision. *Computer Vision and Image Understanding*, 61(1):122–137, January 1995.

[Yi90] Seungku Yi. *Illumination Control Expert for Machine Vision: A Goal Driven Approach*. PhD thesis, Univ. of Washington, 1990.

[YT99] Yiming Ye and John K. Tsotsos. Sensor planning for 3D object search. *Computer Vision and Image Understanding*, 73(2):145–168, February 1999.

[Yua93] Xiaobu Yuan. *3D Reconstruction as an Automatic Modeling System*. PhD thesis, Univ. of Alberta, 1993.

[Yua95] Xiaobu Yuan. A mechanism of automatic 3D object modeling. *IEEE Trans. PAMI*, 17(3):307–311, March 1995.

[ZMH98] Hongbin Zha, Ken'ichi Morooka, and Tsutomu Hasegawa. Next best viewpoint (NBV) planning for active object modeling based on a learning-by-showing approach. In *Third Asian Conf., on Computer Vision, Hong Kong*, pages 185–192, January 8-10 1998.

[ZMHN97] Hongbin Zha, Ken'ichi Morooka, Tsutomu Hasegawa, and Tadashi Nagata. Active modeling of 3D objects: Planning on the next best pose (NBP) for acquiring range images. In *Int. Conf. on Recent Advances in 3D Digital Imaging and Modeling, Ottawa*, pages 68–75, May 12-15 1997.