

View Planning for Multi-Stage Object Reconstruction

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

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

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

Abstract

Previous work has presented a view planning concept for automated object reconstruction. The multi-stage concept comprises a rapid initial exploration stage to produce a low resolution approximate object model - the rough model. With the rough model as the new knowledge base, a view plan is developed for a subsequent precision measurement stage capturing detailed object geometry in accordance with a pre-defined model specification. This paper addresses the issue - “How rough can the rough model be for effective view planning?”. We present experimental results showing that good results are achievable for a reasonable level of rough model decimation, surface errors well within expected levels and sparse sampling of viewpoint space.

1 Introduction

The list of industrial, cultural and other applications requiring high quality 3D object models grows steadily, yet the model acquisition process remains labour intensive and error prone, requiring imaging specialists with extensive training and experience. The conventional process involves an iterative cycle of view planning, sensing, registration and integration. The major remaining hurdle to automated model acquisition is an efficient, accurate and robust solution to the view planning component of this cycle.

The reconstruction process commences with limited a priori knowledge of the object and its immediate environment. Consequently, there is a requirement for both scene exploration and precision measurement, tasks which need not use the same technology. Most current view planning techniques combine these activities in iterative exploration/measurement steps using a single sensor. Each step exploits the current state of knowledge for cues to the next-best-view (NBV). The most common surface-based view planning approach exploits occlusion edges, for example [6], while solid geometry algorithms have been used [10] to model volumetric object knowledge. Other approaches [7] rely on global rather than local characteristics of the geometric data. For a review of these non-model-based

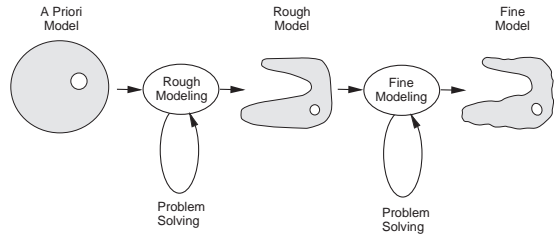


Figure 1: Multi-stage Model-based View Planning

techniques, see [11], [8].

Few model-based approaches to the view planning problem (VPP) have appeared in the literature. An exception is Tarbox and Gottschlich [12] who introduced a model-based concept for inspection using measurability matrices. Subsequently, Scott et al [8] introduced a concept, the Modified Measurability Matrix (3M) algorithm, for performance-oriented view planning for automated model acquisition. *Performance-oriented* reconstruction was 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 requirements are specified in terms of precision and sampling density. This is presently the only known view planning approach to incorporate specific objective reconstruction criteria beyond all-aspect coverage.

Illustrated at Figure 1, the concept 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. The rough model is then used to develop a view plan for a subsequent stage of precision measurement in compliance with the model specification.

As an additional measure to reduce computational complexity, the rough model may be segmented au-

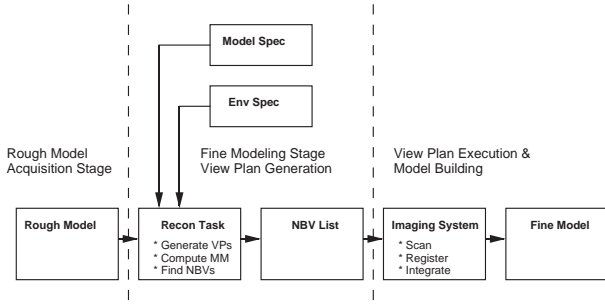


Figure 2: Fine Modeling Process

tomatically or manually into a set of patches based on the view planning challenge presented by different surface features. For example, cavities are particularly difficult to image due to shadow effects. A measurability matrix \mathbf{M} is computed for each segmented region. By convention, rows correspond to surface points and columns to viewpoints.

For inspection, the rough model not only exists but is defined in considerable 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 and yet reliably guide precise 3D geometric sensing and modeling?

2 Experimental Configuration

2.1 View Planning Process

The current work concentrates on the fine modeling reconstruction stage shown in more detail at Figure 2. The process inputs are the rough model plus a model specification and imaging environment specification. The model specification defines reconstruction goals (measurement precision and sampling density) as well as configurable view planning parameters. The environment specification defines key parameters in the sensor and positioning system models. 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 is sent to the physical imaging system to be executed.

2.2 Experimental Process

The experimental procedure followed in this work (Figure 3) employs these steps in a closed loop process with synthetic data to explore the impact of rough model decimation and errors on view plan quality. A

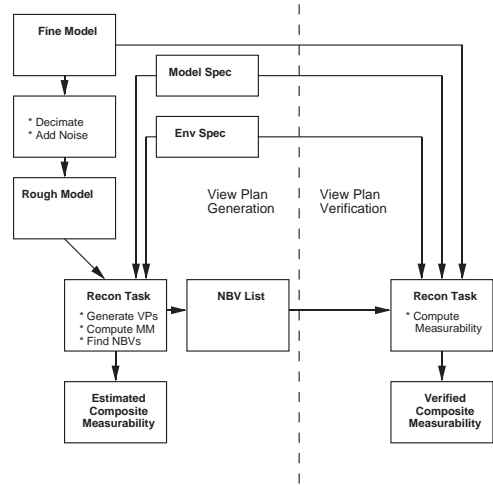


Figure 3: Experimental Process

high density mesh such as the example at Figure 4 serves as a surrogate fine model. The fine model is decimated to create coarser 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 Figure 5. The rough model drives the view planning process, including viewpoint generation, measurability estimation and NBV list determination by means of a greedy search set covering algorithm. Other inputs include imaging environment and model specifications. The resulting view plans are executed against the parent fine model, closing the loop to provide verified measurability.

2.3 Sensor Model

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, scan length, minimum and maximum ranges. The camera is modeled as a bi-static sensor with a specified optical baseline¹. Image size and other parameters are optionally reconfigurable. Viewpoints are treated as generalized viewpoints (\mathbf{v}, λ_s) , consisting of sensor pose \mathbf{v} and a set of controllable sensor parameters λ_s . Thus each viewpoint has an associated sensor configuration. This unrestricted approach to viewpoint generation contrasts with viewing constraints

¹Many conventional view planning approaches, even some of the most sophisticated, treat the range camera as a point sensor and thus fail to take into account one of its most important characteristics - shadow effects.

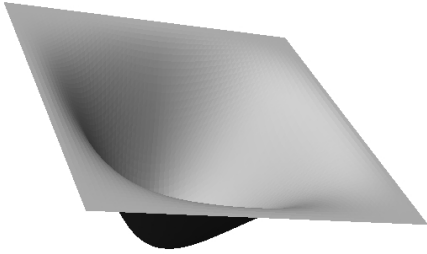


Figure 4: Cavity Fine Model

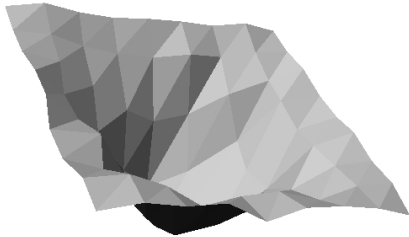


Figure 5: Noisy Cavity Rough Model

commonly applied in the literature, such as the use of “view spheres”. While our positioning system model incorporates error parameters, it is considered error-free for the experiments presented here.

Most importantly, a sensor geometric noise model is incorporated based on range camera calibration data [2], [3], [4]. 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. In this first order model, off-diagonal noise covariance matrix elements are considered secondary effects ignored for view planning purposes.

$$\begin{aligned}\hat{\sigma}_x &= C_x z \\ \hat{\sigma}_y &= C_y z \\ \hat{\sigma}_z &= C_z z^2\end{aligned}\quad (1)$$

As noise along the sensor boresight predominates and small errors tangential to the surface are of debat-

able 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. Thus, we modify the estimated precision obtained from Equation 1 by the following, where $U_i(\theta)$ is the inverse unit step function, ($U_i(\theta) = 1 - U(\theta)$):

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

While grazing angle effects are frequently modeled in view planning algorithms ([1], [5], [12]), they have been treated subjectively rather than objectively. That is, low grazing angle viewpoints are given preference over high grazing angles, rather than evaluating the objective effects of grazing angle on measurement quality, such as precision and sampling density.

2.4 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 s_i on the rough model from a single viewpoint \mathbf{v}_j [8]. 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. Specifically, the surface point must fall within the sensor frustum for that viewpoint, the surface point must be locally visible to the optical source and receiver, and estimated measurement precision and sampling density must meet the specified requirement. Our measurability matrices are binary. Analogue forms can also be defined.

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.

$$m(\mathbf{v}_j) = \frac{|S_j|}{|S|} \quad (3)$$

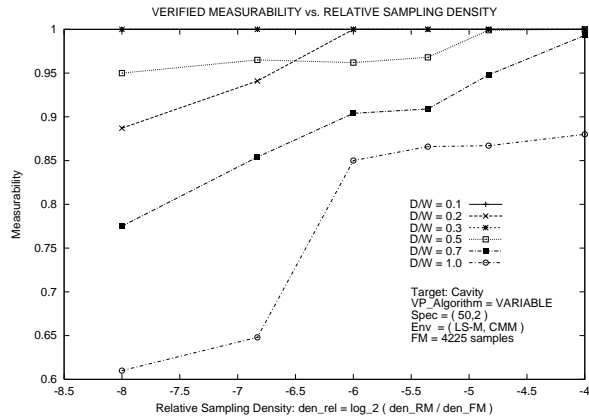


Figure 6: Verified Measurability vs. Relative Sampling Density

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.

2.5 Experiments

Experiments were conducted with synthetic cavity patches of various shapes as representative of a difficult view planning challenge. 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 geometric surface noise and the shape of the target surface patch. For output performance measures, we examine composite measurability 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.

3 Geometry and Relative Sampling Density

Shadow effects in surface concavities present a major scanning problem and view planning challenge when

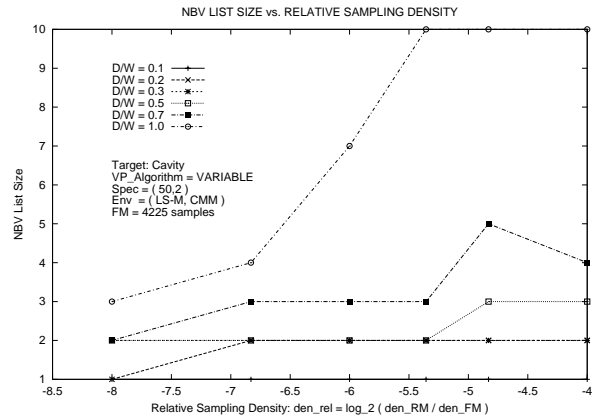


Figure 7: NBV List Size vs. Relative Sampling Density

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. In this section we examine the impact of target shape and relative sampling density on view planning performance.

Experiments were conducted with a variety of cavity shapes and rough model approximations. Figure 6 plots verified measurability as a function of rough model to fine model relative sampling density² ρ_{rel} for a cavity target over a range of geometry expressed in terms of depth-to-width ratio D/W . In general, we would anticipate encountering 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. In our algorithm, discretization levels in viewpoint and object surface space are related. 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

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

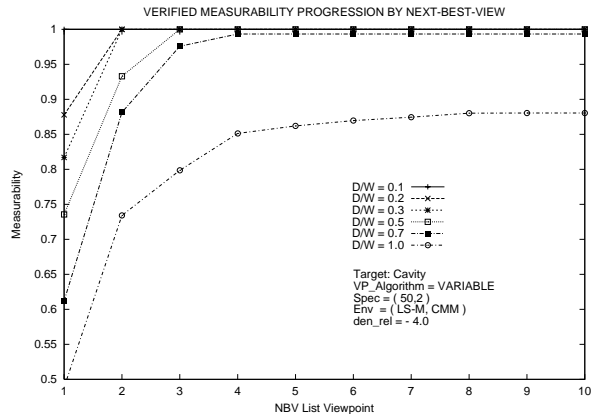


Figure 8: Verified Measurability Progression with NBV Selection

around $D/W \approx 1.0$.

The second key performance indicator is the size of the NBV list. An approximate indicator of the number of views required for a patch [9] is provided by $\lceil r_f \rceil$ where the footprint ratio $r_f = F_p/F_s$, the patch footprint F_p is the patch area and sensor footprint F_s is the frustum cross-sectional area at the sensor’s optimal scanning range. 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 7 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, the NBV list size was capped at ten views.

Figure 8 shows the progress of verified cumulative measurability with successive views from the NBV list for one rough model size. Viewpoints generated from the rough model are selected by a greedy set covering algorithm on the basis of declining utility. Actual scanning results may vary slightly. 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.

An Extended Gaussian Image (EGI) view of the distribution of surface normals for the very deep mesh ($D/W = 1.0$) used in the experiment at Figure 8 is

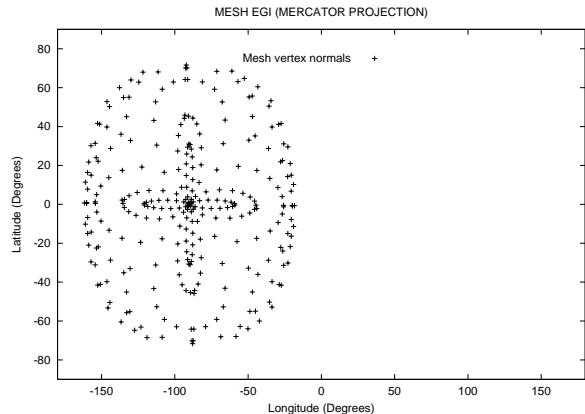


Figure 9: Mesh Extended Gaussian Image

shown at Figure 9 in Mercator projection format³. 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 10 presents an EGI view of the axis component of the NBV viewpoint list generated for this target. As expected, NBV list viewpoints are situated on the opposite side of the EGI sphere to the target surface. Both the NBV list and its negative mirror image are displayed for 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 summary, the relative sampling density level ρ_{rel} determines discretization intervals in both rough

³An 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 Greenwich axes can be selected to minimize distortion.

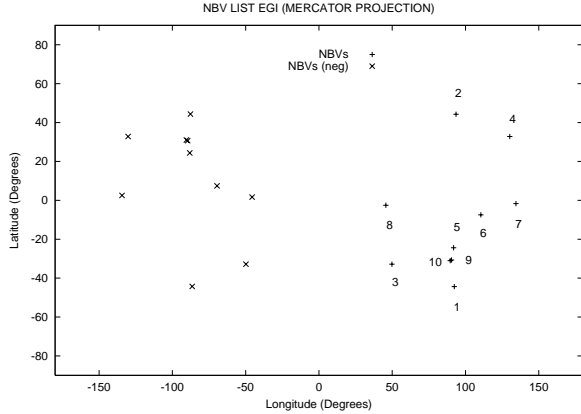


Figure 10: NBV List Extended Gaussian Image

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 $\rho_{rel} \approx -5$ 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. Further, for typical surface patches and imaging geometries, most scanning coverage and measurement is achieved with a very small number of scans (views).

4 Rough Model Errors

By definition, the 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 surface errors on view planning performance.

View planning algorithms favour measurement geometry along the local surface normal. The geometric noise component along the sensor boresight greatly exceeds transverse measurement error. Hence we select a simple model for rough model surface error based on a noise component along the local surface normal. The modeled noise is a zero mean process with a standard deviation σ_{rm} expressed as a percentage of the sensor's minimum range. In the present work we do

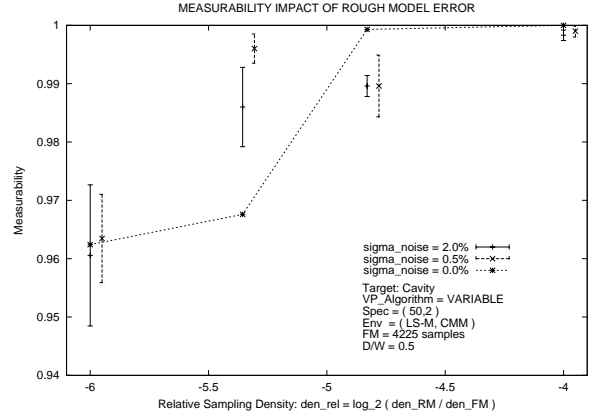


Figure 11: Measurability Impact of Rough Model Error

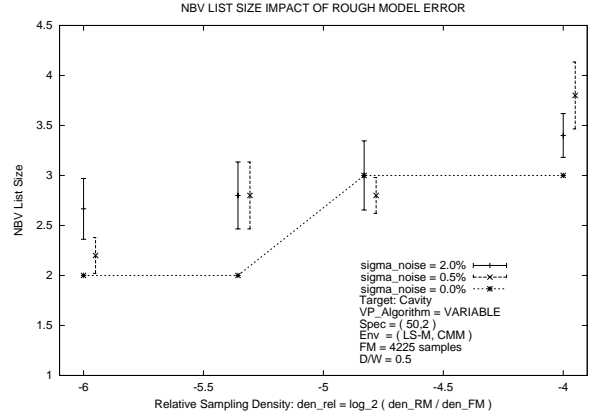


Figure 12: NBV List Size Impact of Rough Model Error

not address outliers, edge effects, corner reflections or 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 (Figure 5). Figures 11 and 12 show the impact on verified measurability and the size of the NBV list of a noisy rough model in comparison to the noiseless case⁴. The noisy data is presented as average

⁴Note that the data for $\sigma_{rm} = 0.005R_{min}$ is artificially separated slightly to the right for readability.

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 essentially 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.

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 error sources from a more extensive covering of viewpoint space.

5 Summary and Conclusions

It is instructive to examine actual versus estimated measurability performance. Figure 13 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 14 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 15. 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 $\mathcal{O}(vs^2)$, where $v = |V|$ and $s = |S|$ are the sizes of the viewpoint and surface point sets, respectively. A key feature of the algorithm is sparse sampling in surface and

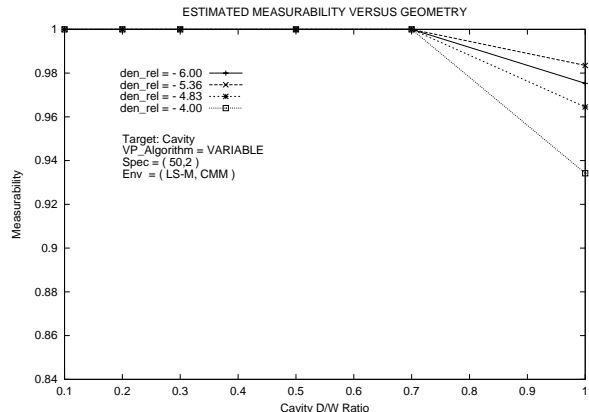


Figure 13: Estimated Measurability Vs. Cavity Geometry

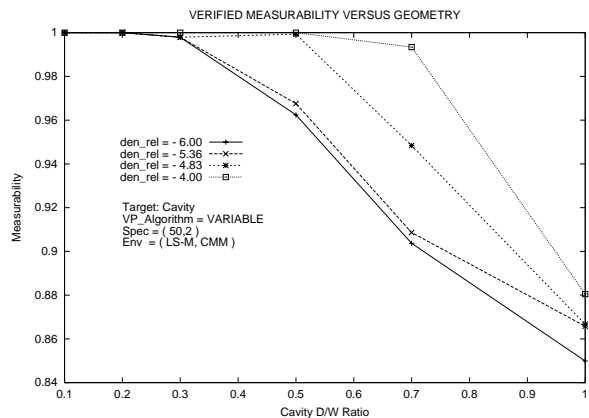


Figure 14: Verified Measurability Vs. Cavity Geometry

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 s of the rough model segment also influences the size v of the corresponding viewpoint set. The experimental results of this section indicate that good results are achievable for a level of rough model decimation well within the desired range. We have also shown that the algorithm is robust with respect to fairly high levels of rough model surface error.

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

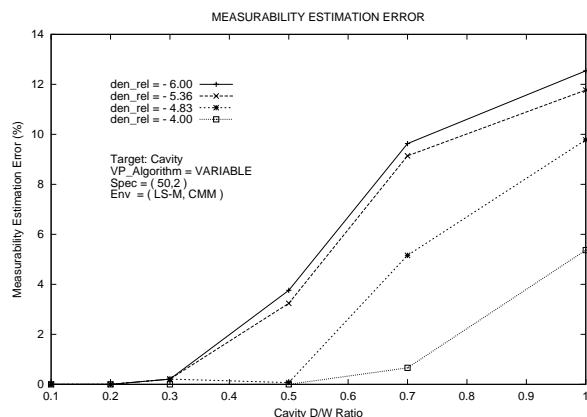


Figure 15: Measurability Estimation Error Vs. Cavity Geometry

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 2.3 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.

In conclusion, 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 well within expected levels and a sparse sampling of viewpoint space. Future work will enhance the sensor model to handle dynamic range limitations.

References

- [1] J. E. Banta and M. A. Abidi. The positioning of a range sensor for optimal reconstruction of three-dimensional models. In *Int. Conf. on Recent Advances in 3-D Digital Imaging and Modeling, Ottawa, May 12-15 1997*.
- [2] J.-A. Beraldin, S. EL-Halim, and L. Cournoyer. Practical range camera calibration. In *Proc. Videometrics II, SPIE, Boston, Massachusetts*, volume 2067, pages 21–30, 9-10 September 1993.
- [3] S. EL-Halim 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.
- [4] S. EL-Halim 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.
- [5] N. A. Massios and R. B. Fisher. A best next view selection algorithm incorporating a quality criterion. In *British Machine Vision Conference 1998*, pages 780–789, September, 1998.
- [6] J. Maver and R. Bajcsy. Occlusions as a guide for planning the next view. *IEEE Trans. PAMI*, 17(5):417–433, May 1993.
- [7] R. Pito. A sensor based solution to the next-best-view problem. In *IEEE Int. Conf. on Robotics and Automation*, pages 941–945, August 1996.
- [8] W. Scott, G. Roth, and J.-F. Rivest. Performance-oriented view planning for automatic model acquisition. In *31st Int. Symposium on Robotics, Montreal*, pages 314–319, May 2000.
- [9] W. Scott, G. Roth, and J.-F. Rivest. View planning with a registration constraint. In *Submitted to 3rd Int. Conf. on 3-D Digital Imaging and Modeling, Quebec City, Canada, 2001*.
- [10] I. Stamos and P. Allen. Interactive sensor planning. In *Proc. IEEE Conf. Vis. Pat. Rec., Santa Barbara, CA*, pages 489–494, June 23-25 1998.
- [11] K. Tarabanis, P. K. Allen, and R. Y. Tsai. A survey of sensor planning in computer vision. *IEEE Trans. Robotics and Automation*, 11(1):86–104, February 1995.
- [12] G. Tarbox and S. Gottschlich. Planning for complete sensor coverage in inspection. *Computer Vision and Image Understanding*, 61(1):84–111, January 1995.