

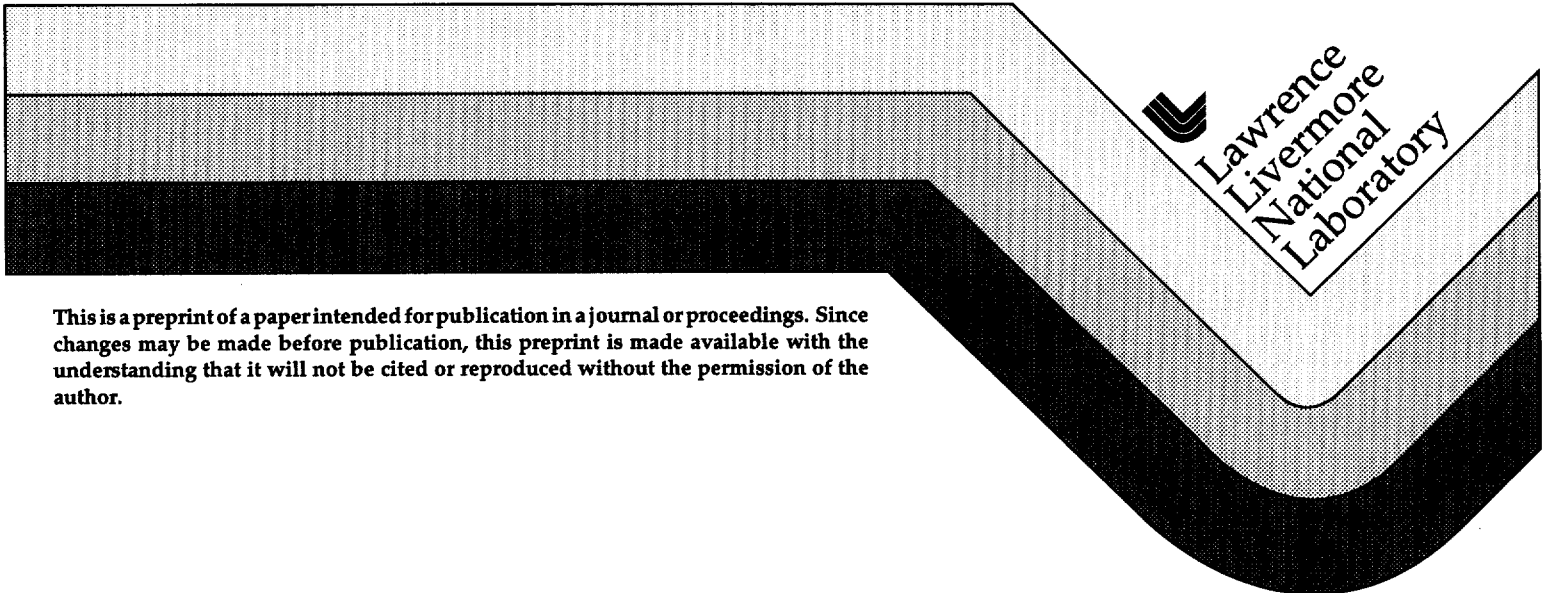
UCRL-JC-128921
PREPRINT

A Human Supervisory Approach to Modeling Industrial Scenes Using Geometric Primitives

J. P. Luck
C. Q. Little
R. S. Roberts

This paper was prepared for submittal to the
International Conference on Robotics & Automation '98
Leuven, Belgium
May 16-21, 1998

November 19, 1997



This is a preprint of a paper intended for publication in a journal or proceedings. Since changes may be made before publication, this preprint is made available with the understanding that it will not be cited or reproduced without the permission of the author.

DISCLAIMER

This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor the University of California nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or the University of California, and shall not be used for advertising or product endorsement purposes.

A Human Supervisory Approach to Modeling Industrial Scenes Using Geometric Primitives*

Jason P. Luck¹

Colorado School of Mines
Golden, CO 80401 USA

Charles Q. Little

Sandia National Laboratory
Albuquerque, NM 87185 USA

Randy S. Roberts^{1,2}

Lawrence Livermore National Laboratory
Livermore, CA 94550 USA

ABSTRACT

A three-dimensional world model is crucial for many robotic tasks. Modeling techniques tend to be either fully manual or autonomous. Manual methods are extremely time consuming but also highly accurate and flexible. Autonomous techniques are fast but inflexible and, with real-world data, often inaccurate. The method presented in this paper combines the two, yielding a highly efficient, flexible, and accurate mapping tool. The segmentation and modeling algorithms that compose the method are specifically designed for industrial environments, and are described in detail. A mapping system based on these algorithms has been designed. It enables a human supervisor to quickly construct a fully defined world model from unfiltered and unsegmented real-world range imagery. Examples of how industrial scenes are modeled with the mapping system are provided.

1.0 INTRODUCTION

There are many uses for intelligent, semi-autonomous robots in industrial environments. In particular, industrial environments such as contaminated processing plants or hazardous waste sites can benefit from intelligent robots during decontamination and decommissioning operations. An important ingredient of semi-autonomous robotics is a world model of the environment in which the robot operates. An accurate world model allows the robot to plan and execute its tasks in a robust fashion. Unfortunately, world models for many industrial sites are either unreliable or unavailable. In these situations, the ability to construct an accurate world model is a necessity for robot deployment [1].

The traditional method of generating world models is to manually create the models using CAD tools. This method is not only extremely time consuming but may

be impossible in hazardous areas in which object measurements cannot be collected. Fully autonomous methods have also been tested in recent years, but these methods are not reliable enough to be used with real-world data on hazardous tasks [2]. Many papers model scenes with geometric primitives (see [3] and [4]). Several of these approaches can find the position and orientation of an object, but they need a predetermined model. Others adeptly create models from segmented data but do not deal with obscured regions, object intersections, or approach the problem of segmentation. Many methods use edge detection to create well-defined object intersections and to aid segmentation, such as [5]. These methods work well on complete data sets but have problems when edges are obscured. Because of the problems associated with data of this type, the approach presented in this paper was developed.

The techniques presented in this paper are designed to build accurate world models from multiple range images collected from industrial environments (see [6] and [7]). The data used to develop the algorithms were collected from a structured lighting system at Sandia National Laboratories. The data is noisy and contains errors that render most segmentation and modeling algorithms inadequate. A high concentration of reflectance points (artifacts of the structured lighting system) permeate the data set and cause large variations in surface normal and curvature values for all points surrounding the reflectance point. Other errors, such as misregistration of multiple images and additive noise, also cause deviations in the surface normal and curvature. These types of problems result in the failure of segmentation and modeling algorithms that exploit characteristics of the local surface geometry. Additionally, the range samples are non-uniformly spaced, meaning the distance between two neighboring points on the x-sampling axis can be much larger than the distance between two neighboring points on the y-sampling axis. This non-uniformity causes preferential growth in the direction of higher point density and failure of many region-growing algorithms. Registration errors cause large abnormalities when multiple images of the same scene are overlaid. Moreover, there are often large obscured regions in the imagery, which cause undefined intersections between planes and holes in the model.

*This work was partially supported by the Defense Advanced Research Projects Agency, and partially supported by the Department of Energy.

¹Formerly with Los Alamos National Laboratory, Group ESA-EPE.

²Corresponding author.

Because of the large extent of errors in real-world data and the lack of reliability in autonomous methods, we have opted to use a human supervisor to guide the modeling process. The underlying approach is to use geometric primitives to fuse and segment the range imagery, and then use the segmented data to refine the parameters of the geometric primitive. Geometric primitives are ideal for modeling industrial environments since they require very little memory for large world maps, and are easy to store and manipulate. Once the parameters for all the surfaces in the scene are found and refined, the geometric models are extrapolated to build the final world model. The world model is then exported to a CAD system.

From the supervisor's point of view, the modeling system consists of two windows and a few controls on a workstation. The first window displays range imagery from the structured lighting system, and the second window contains the world model built from the range imagery. The range imagery appears in the window as a collection of points in space. The supervisor selects initialization points on the surface of an object, and instructs the computer to fuse, segment and model the object. As an object is modeled, the data associated with the object is highlighted in the data window. When the operator is satisfied with the fusion and segmentation, the model is refined and the surface represented by the model appears in the model window. The points associated with the object in the data window are also removed, allowing the supervisor to easily monitor progress. After the range data has been fused, segmented and modeled, the operator extrapolates surfaces in the model window until a fully defined world model is produced.

With regards to geometric primitives used to model the surfaces, casual observation of industrial environments reveals that a few types of surfaces can be used to model a wide range of industrial scenes. Bounded planes, for example, can be used to model walls, ceilings and floors, and combined to model cabinets, desks, crates and other items. Finite-length cylinders can be used to model ventilation ducts, chemical storage drums, electrical conduit among other objects. A final type of surface, a random surface, can be used to model objects such as sand heaps or a pile of debris. The details of how we fuse, segment and model these surfaces are described next.

2.0 DETAILED DESCRIPTION

Let $S_i = \{(x_k, y_k, z_k)\}$ be the i^{th} set of range samples collected from surfaces in an environment. The samples S_i are typically collected from a ranging system (e.g., structured lighting) and provide a single view of the environment. Consider the union of N sets of range samples of the same scene but collected from different viewpoints. The set of range samples, $S = S_1 \cup S_2 \dots \cup S_N$, thus represent samples of surfaces in the environment. Our objective is to build a reliable scene model from the data in S .

2.1 Planes

The supervisor typically starts by fusing, segmenting and modeling the planes in S . A bounded plane is described by a set of parameters $p = \{a, b, c, d\}$ which define the (infinite) plane $ax + by + cz = d$, and a set of edge points $e = \{(x_k, y_k, z_k)\}$, that lie in the plane and represent the vertices of the plane's boundary. The problem is now one of determining plane parameters p and edge points e . An estimate of p is found as the least-squares solution to a system of equations formed from a set of initialization points $u = \{(x_k, y_k, z_k)\}$, selected by the supervisor. Typically, the supervisor randomly selects four to ten initialization points scattered around the plane and its periphery. After parameters p have been estimated, a new (primed) coordinate system is established such that the $z' = 0$ plane of the new coordinate system coincides with the plane represented by the estimated parameters in p . The data in S is then rotated from the world coordinate system into the primed coordinate system and denoted as S' . The standard deviation $\sigma_{z'}(u')$ of the perpendicular distance of the rotated initialization points to the plane $z' = 0$ is now estimated. A new set of points, $T' \subseteq S'$ is derived from S' such that

$$T' = \{(x'_k, y'_k, z'_k) \mid |z'_k| \leq \lambda \sigma_{z'}(u')\}$$

where λ is a tuning parameter used to fuse misregistered planes. By correctly adjusting λ , data from misregistered planes can be assimilated into a single plane. Effectively, a one-dimensional (i.e., along the z' coordinate) segmentation of the data in S' is performed. The next task is to segment the data in T' along the x' and y' coordinates.

Two dimensional segmentation of the remaining points in T' begins by forming a bounding box B of width $\Delta x'$ and height $\Delta y'$ around the initialization points $u'_0 = \{(x'_k, y'_k, 0)\}$. The width and height of the bounding box is then expanded to $2\Delta x'$ and $2\Delta y'$ respectively, and points in T' that are exterior to the expanded bounding box are discarded. This simple operation is adequate for isolated planes, e.g., the top of a box, but inadequate for planes touching other surfaces, e.g., the side of a box sitting on a floor. Thus, additional steps are performed on the data in T' to correctly segment planes such as the latter. To start with, all remaining points in T' are projected onto the x' axis, and a ten-bin histogram over the region of support of the distribution is calculated. If a bin contains less than 15% of the peak value of the histogram, the points in that bin and all points from that bin to the current plane boundary are discarded. The same process then occurs on the y' axis. If points are discarded, the histo-

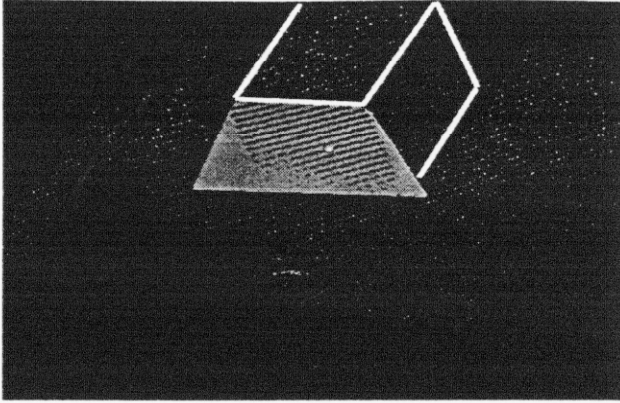


Figure 1: Illustration of wing artifacts on the front face of a box on the floor. The image consists of three overlaid scans of the box, and portions of the box have been outlined for clarity. The shaded region is the boundary of the box with wing points included in the boundary computation.

gram process is repeated until no new points are eliminated.

One artifact that can be introduced by this segmentation technique is *wing points*. Wing points are outliers that are not removed by the histogram operations, and form wings on the plane when its boundary is estimated. Figure 1 illustrates wings on the side of a box caused by wing points on the floor. In general, wing points (if they are present) occur near one boundary of the plane. They are eliminated by a procedure that forms an upper boundary $x' = a$, and initializes to $a = \bar{x}'$, where \bar{x}' is the mean abscissa value of points in T' . Define two subsets of T' :

$$T'_1 = \{T' | x'_k \geq a, y'_k \leq \bar{y}'\}$$

$$T'_2 = \{T' | x'_k \geq a, y'_k \geq \bar{y}'\}$$

where \bar{y}' is the mean ordinate value of the points in T' . The value of a is increased to the point where either $T'_1 = \emptyset$, or $T'_2 = \emptyset$. Similar operations are performed for the lower boundary in x' , and the y' boundaries to eliminate any wing points.

After the data in T' has been segmented, it is rotated back into the world coordinate system and the parameters of the plane are refined (i.e., recomputed) using *all* of the remaining data in T . Denote the final (optimal) parameters of the plane as $p_o = \{a_o, b_o, c_o, d_o\}$. A second rotation is performed such that plane described by p_o coincides with the plane $z' = 0$. The convex hull of the set of points $\{(x'_k, y'_k, 0)\}$ is next computed. The points in the convex hull form the boundary of the plane in the rotated coordinate system, e' . Finally, the edge points in the world coordinate system are found by back-rotating

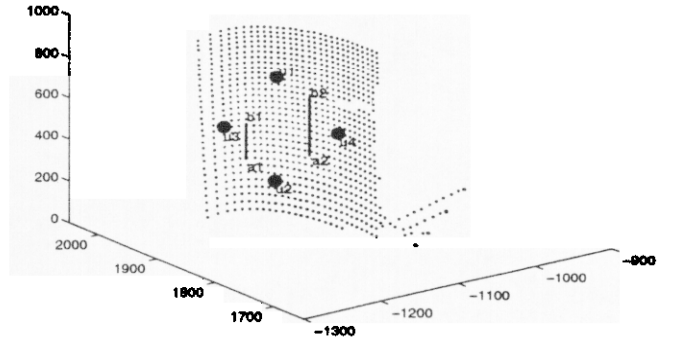


Figure 2: Estimating the axis of a cylinder. The range image of a part of a cylinder is shown, concave side towards the viewer. The four points labeled u_k , $k = 1, \dots, 4$, are the operator selected initialization points. The vertical lines are the result of the axis search.

e' . To simplify succeeding fusion and segmentation operations, the data points in T are removed from S .

2.2 Cylinders

A bounded cylinder is defined by four parameters: a scalar radius, r ; an axis vector, a ; a center point to fix the axis vector, $c = (x_c, y_c, z_c)$; and a set of points $h = \{(x_k, y_k, z_k)\}$ that represents the boundary (height) of the cylinder. Processing begins by estimating the axis vector of the cylinder using a modification of the technique described in [8]. Four seed points, $u = \{(x_k, y_k, z_k)\}$, chosen with the first two pointing approximately along the axis, are used to form a region on the side of the cylinder (see Figure 2). The region is divided into halves, and two parallel vectors are sought with one in each half. The result of the search is shown in Figure 2, where two vectors extend from points a_1 to b_1 and from points a_2 to b_2 . Note that if two parallel vectors are found on the surface of a cylinder, they must also be parallel to the cylinder axis. The estimate of the cylinder's axis, \hat{a} , is used to construct a rotation matrix, and data in S is then rotated into the primed coordinate system S' so that the cylinder axis is parallel to the z' axis. Projecting data in S' onto the $z' = 0$ plane results in an arc of samples for an incomplete cylinder, or an entire circle for a complete cylinder (see Figure 3). Now that the data has been projected onto the $z' = 0$ plane, parameters r and c can be estimated.

Recall that the center of a circle can be found at the intersection of two perpendicular bisectors of any two

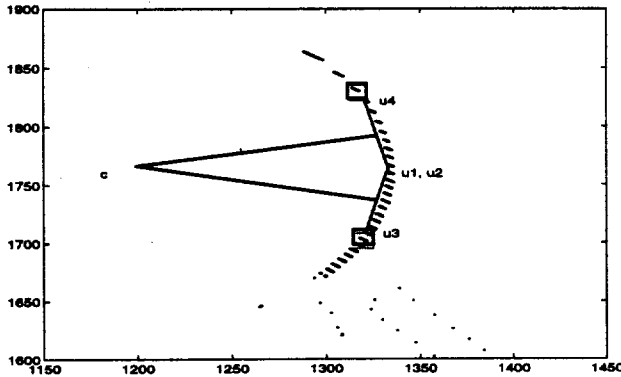


Figure 3: Estimating the center and radius of a cylinder. The cylinder is rotated so that the main axis is parallel to the z' axis, and the data projected into the $z' = 0$ plane. From there, the center and radius are estimated.

chords of the circle. The radius of the circle is then found as the distance from the center of the circle to any point on the circle. The transformed initialization points u' are known to lie within the arc of points formed by projecting the cylinder onto the $z' = 0$ plane. These points are connected into chords in the circle and used to estimate the center of the circle \hat{c} . An initial estimate of the radius, \hat{r} , is found as $\hat{r} = \text{mean}(|\hat{c} - u'|)$. Estimates \hat{c} and \hat{r} can be improved by using larger chords, which are found by iteratively moving the ends of the chords, u_3' and u_4' , to the ends of the arc. Small neighborhoods (shown as boxes) surrounding u_3' and u_4' are first defined by the seed point itself, a point using the same radius but 7.5 degrees closer to the end of the arc, and by a factor times the deviation of the radius of all points within the previous neighborhood σ_r (for the first iteration σ_r is set to 1). If points are found within the neighborhood bounds, either u_3' or u_4' is redefined at the center of the neighborhood bounds and both \hat{c} and \hat{r} are re-estimated. Continued iterations provide refined estimates of \hat{c} and \hat{r} , and also an estimate of the variance of the radius estimate σ_r . The variance estimate, together with \hat{c} and \hat{r} , allow samples of the cylinder to be segmented from the rest of the samples in S' . These segmented samples form a set T' , where $T' \subseteq S'$.

The final values of a , c , and r are provided by the Levenberg-Marquardt non-linear least-squares algorithm operating on data in T' . The previously obtained estimates of \hat{r} , \hat{c} , and \hat{a} are to initialize the algorithm. The algorithm minimizes the deviation of each data point by adjusting the model parameters using the partial derivatives of each parameter. The process iterates itself until it converges upon the best set of parameters.

Several different values of cylinder height are possible. For example, the height of a right cylinder can be described by a single number, whereas a cylinder with

arbitrary extent requires a complex description of its height. Anticipating complex cylinders, we describe the cylinder's height as a set of coordinates h . Data that forms the cylinder is rotated and translated so that the z' axis is parallel to the cylinder axis, and the origin is located at the center of the cylinder. The cylinder is segmented into 10-degree sectors, and the maximum and minimum z' value is found for each sector. These values form the set h' , which when rotated back into the world coordinate system describe the extent of the cylinder.

2.3 Random Surfaces

Random surfaces are modeled as a connected group of triangles. To begin, the supervisor creates a plane as described in Section 2.1, and sets the fusion parameter λ large enough to contain the random surface. The data is rotated and projected onto the plane. A standard Voronoi tessellation routine [9] creates a connectivity list, ignoring the z' component. The z' component is then reintroduced into the data, and the data is then rotated back to the world coordinate system, where triangles are drawn between the points according to the connectivity list.

3.0 MODEL BUILDING

Construction of the world map is performed after the surfaces in the scene have been fused, segmented and modeled. Say that the supervisor found N_P planes, N_C cylinders and N_R random surfaces in the scene. Since these surfaces are stored as equations, parameters and bounding points, extrapolation of the surfaces beyond the estimated boundaries, and interpolation of the surfaces into shadows is relatively easy. The supervisor selects two objects to be joined, and instructs the modeling system to join them. If the objects are non-parallel planes, the line of the intersection is calculated, and new convex hulls for each plane are calculated. The new convex hulls include the points along the intersection, and any points extending past the intersection are eliminated. This method is also extended to encompass three adjacent planes, such as the corner of a room.

Boxes are so common in industrial environments that they are considered a special case. As a result of occlusion, boxes usually appear as only three planes. The other sides are not visible to the sensors, but three sides are sufficient to reconstruct the box. The assumption that a side is parallel to (and the same size as) the opposing side allows the reproduction all six sides of a box from three adjacent planes. First, the equations and convex hulls of each of the three adjacent planes are computed. A coordinate system with the plane intersection point as the origin and with each axis pointing along a plane intersection line is next formed. The dimensions of the box are equal to the maximum point value in each axis direction. Using the previ-

ously stated assumptions, the routine creates points at the corners of the box and draws planes connecting the corner points.

If the object is a cylinder pointing towards a plane, it too can be modified so that the cylinder extends to the plane. Because the equations of both the cylinder and plane are known, the set of points representing the height of the cylinder h can be extended or drawn back until they lay in the plane. Similarly, if a plane covers a portion of a cylinder, such as the lid of a drum, the plane can be extended until it meets the cylinder boundary. The convex hull of the extended plane is then computed and stored.

The process of intersecting objects and building boxes eventually yields a fully defined world model. Denote the final planes as P_i , the final cylinders as C_j , and the final random surfaces as R_k . The world model M is then given as the union of all of these objects:

$$M = \bigcup_{i=1}^{N_p} P_i \cup \bigcup_{j=1}^{N_c} C_j \cup \bigcup_{k=1}^{N_r} R_k$$

As a final step, world model M is exported to a CAD system for use by a robot.

4.0 EXAMPLES

Two examples are provided that illustrate the mapping method described in this paper. Range imagery for the examples was collected on a structured lighting system at Sandia National Laboratory, and contains real-world imperfections such as misregistration, shadows and noise. The first example consists of two drums on the floor near the corner of a room (see Figures 4-7). The image in Figure 4 is a result of overlaying two range images of the scene collected from different viewpoints. Note the range shadows on the floor and wall. Although it is not apparent in Figure 4, there is significant misregistration between the two range images on the plane of the floor. Figure 5 shows the five initialization points used to fuse and segment the plane of the floor. The shaded region in Figure 5 is the data that has been associated with the plane, and provides visual feedback to the supervisor on the fusion and segmentation operation. Figure 5 also shows the four initialization points used to segment the larger drum, and a wire-frame model of the drum built from the segmented data. Like the shaded regions for planes, the wire-frame cylinder provides visual feedback to the supervisor on modeling progress. Figure 6 shows the three disconnected planes, and two cylinders that result from fusion and segmentation operations. Figure 7 illustrates the final world model in which the planes have been joined to form complete walls, a floor, and a corner, and the cylinders have been extrapolated so that they touch the floor.

The second example illustrates modeling a random surface (see Figures 8-10). Figure 8 is a single range

image of a vermiculite pit containing several plates and a cylinder. The two regions exterior to the pit are portions of a floor, and are not of interest to the world model. Figure 9 shows the initialization points used to define a bounding box around the random surface. The height of the bounding box has been adjusted so that all of the random surface is enclosed within the box. (Note that the plates and cylinder have already been segmented out of the image in Figure 9.) Figure 10 shows the final world model of the scene.

5.0 BIBLIOGRAPHY

- [1] R. E. Barry, C. Q. Little, and B. L. Burks, "Requirements and Design Concept for a Facility Mapping System," *Proc. ANS 6th Topical Meeting on Robotics and Remote Systems*, Monterey CA, pp.775-783, Feb. 5-10, 1995.
- [2] A. Hoover, G. Jean-Baptiste, X. Jiang, P. Flynn, H. Bunke, D. Goldgof, K. Bowyer, D. Eggert, A. Fitzgibbon, and R. Fisher, "An Experimental Comparison of Range Image Segmentation Algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol 18, no. 7, pp. 673-688, July 1996.
- [3] W. A. Hoff, F. W. Hood, and R. H. King, "An Interactive System for Creating Object Models from Range Data based on Simulated Annealing," *Proc. 1997 IEEE Int. Conf. on Robotics and Automation*, Albuquerque, NM, 1997.
- [4] J. Feddema, and C. Little, "Rapid World Modeling: Fitting Range Data to Geometric Primitives," *Proc. 1997 IEEE Int. Conf. on Robotics and Automation*, Albuquerque, NM, 1997.
- [5] J. K. Hackett and M. Shah, "Segmentation Using Intensity and Range Data," *Optical Engineering*, vol. 28, no. 28, June 1989.
- [6] R. S. Roberts, "An Algorithm for Segmenting Range Imagery," Los Alamos National Laboratory Report LA-UR 96-4497, Nov. 1996.
- [7] J. P. Luck and R. S. Roberts, "A Human Supervisory Approach to Rapid World Modeling Through the Use of Geometric Primitives," Los Alamos National Laboratory Report LA-UR 97-3344, Aug. 1997.
- [8] W. E. L. Grimson, T. Lozano-Perez, and D. P. Huttenlocher, *Object Recognition by Computer: The Role of Geometric Constraints*, MIT Press, Cambridge Mass., 1990.
- [9] Steve J. Fortune, "A Sweepline Algorithm for Voronoi Diagrams", *Algorithmica* 2, 153-174, 1987.

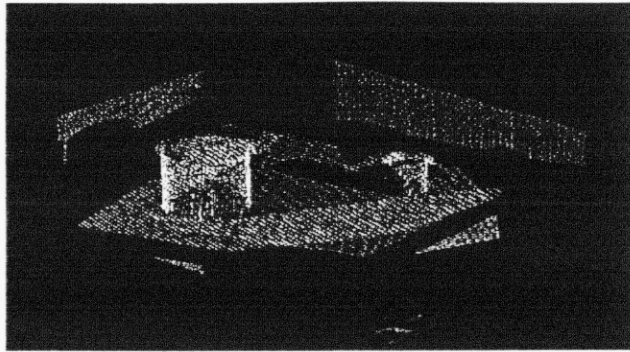


Figure 4: Overlay of two images consisting of two cylinders and several planes.

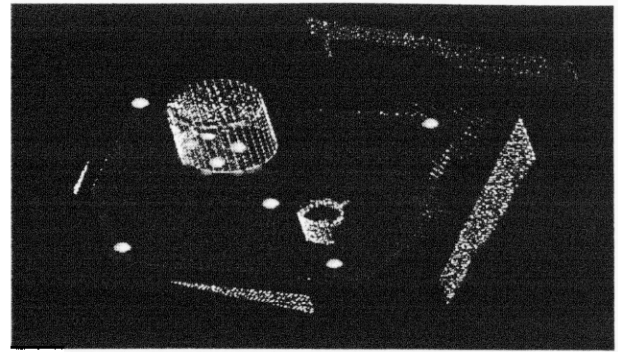


Figure 5: A plane and cylinder are segmented and modeled using the initialization points as shown.

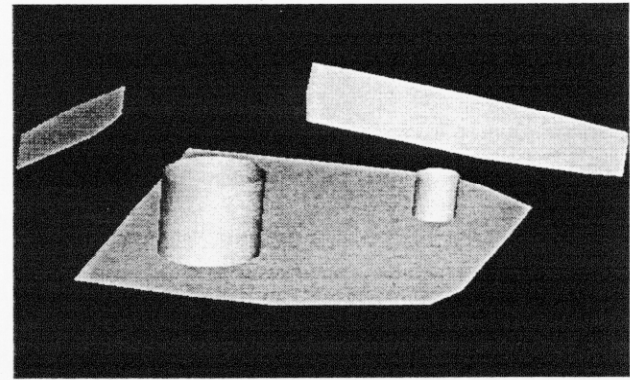


Figure 6: All planes and cylinders have been segmented and modeled.

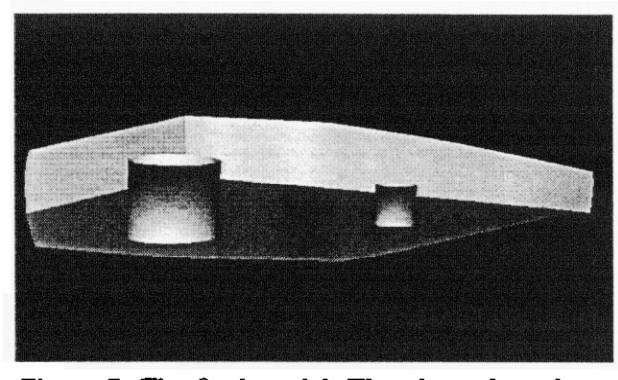


Figure 7: The final model. The planes have been extrapolated to form the corners of the room, and the cylinders extrapolated to rest on the floor.

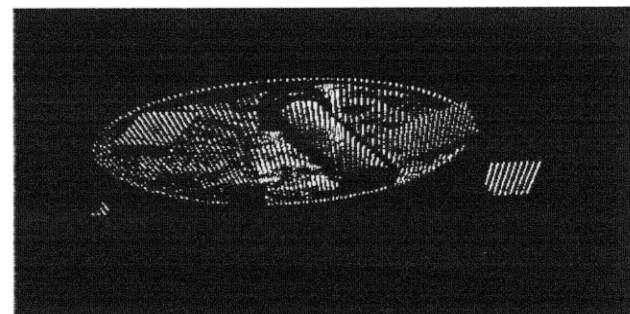


Figure 8: Data set consisting of a pit containing vermiculite, a cylinder and two plates.

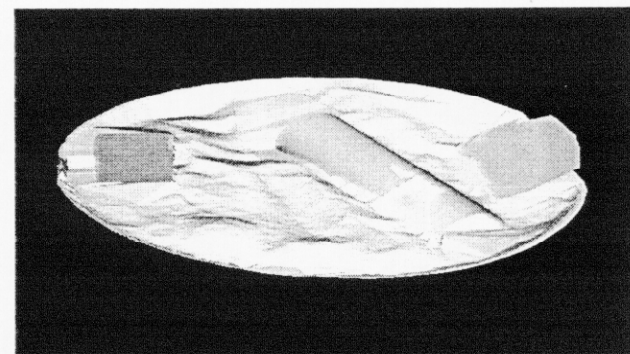


Figure 10: Finished model showing the plates and cylinders integrated into the veronoi tessellation of the vermiculite pit.

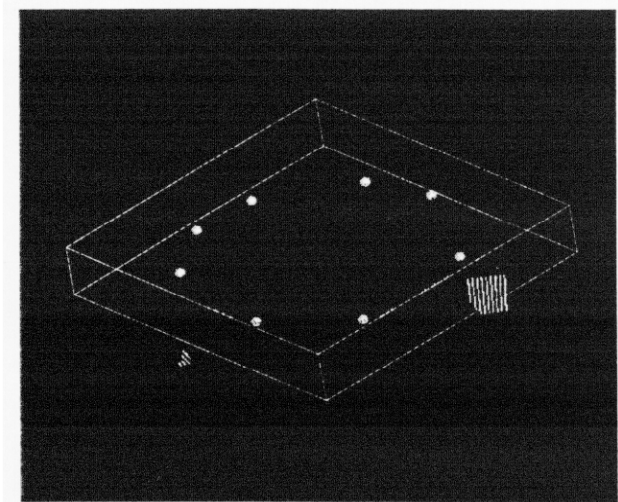


Figure 9: The cylinder and two planes have been modeled and removed. Initialization points used to define a plane are shown, as is the bounding box used to include all points desired for the random surface.

Technical Information Department • Lawrence Livermore National Laboratory
University of California • Livermore, California 94551

