Published in Proceedings of the ANS 7th Topical Meeting on Robotics & Remote Systems, Augusta GA, April 27-May 1, 1997

EVALUATION OF AN INTERACTIVE TECHNIQUE FOR CREATING SITE MODELS FROM RANGE DATA

Frederick W. Hood
Colorado School of Mines
1500 Illinois St.
Golden, CO 80401
(303) 273-3569

William A. Hoff, Ph. D
Colorado School of Mines
1500 Illinois St.
Golden, CO 80401
(303) 273-3761

Robert King, Ph. D
Colorado School of Mines
1500 Illinois St.
Golden, CO 80401
(303) 273-3305

I. ABSTRACT

Recognizing and locating objects is crucial to robotic operations in unstructured environments. To satisfy this need, we have developed an interactive system for creating object models from range data based on simulated annealing and supervisory control[a]. This interactive modeling system maximizes the advantages of both manual and autonomous methods while minimizing their weaknesses. Therefore, it should outperform purely autonomous or manual techniques. We have designed and executed experiments for the purpose of evaluating the performance of our application as compared to an identical but purely manually driven application. These experiments confirmed the following hypotheses:

- Interactive modeling should outperform purely manual modeling in total task time and fitting accuracy.
- Operator effort decreases significantly when utilizing interactive modeling.
- User expertise does not have a significant effect on interactive modeling task time.
- Minimal human interaction will increase performance on “easy” scenes.

Using 14 subjects and 8 synthetically generated scenes, we recorded the task times and pose data and, from them, used analysis of variance (ANOVA) to test a set of hypotheses.

II. INTRODUCTION

The use of robots in hazardous applications such as radioactive facility dismantlement and remediation of buried waste has become a necessity due to the risk to humans in such environments. Control schemes for robots in such an environment range from fully manual (*i.e.*, human teleoperation) to fully automated (*i.e.*, no human supervision or input). Supervisory control measures exist between these two extremes and offer a promising alternative[1]. Systems performing tasks under supervisory control enjoy the flexibility and reliability of human interaction combined with the speed and accuracy of computerized processes[2].

In unstructured environments such as waste dumps where the number, locations, and attributes of objects populating a scene are not known in advance, graphical site models are a necessary tool. These models contain 3-D graphical representations of the objects in a scene and facilitate maneuvering in and manipulation of a scene by robots. The creation of these scenes can be quite tedious, however. Typically, physical measurement of the object locations in the scene is used in conjunction with a graphical user interface or CAD software to produce a site model. These models suffer from many inaccuracies, particularly concerning the orientation of the object since this is difficult to measure accurately by hand. Another obvious consideration is that humans should not take measurements in hazardous environments.

Range data gathered from a stereo vision system, laser range finder, or other sensing devices provides a useful alternative to physical measurement in the creation of site models. Models can be formed from this range data using a wide variety of techniques. In the past, these techniques have been predominately manual or automated. Manual techniques for site modeling often utilize a graphical user interface, with a human operator doing the model placement

[a] This project has been funded by the INEL University Research Consortium. The INEL is managed by Lockheed Martin Idaho Technologies Company for the U.S. Department of Energy, Idaho Operations Office, under Contract No. DE-AC07-94ID13223

and fine-tuning. These methods are historically tedious and slow. Automated techniques, which often employ image understanding algorithms and optimization routines, are a speedy alternative to manual techniques. However, purely automated systems lack the reliability of purely manual systems. Therefore, we have developed an interactive system for fitting object models to range data. An interactive system consists of both manual and automated components and utilizes their respective advantages. In this paper, we briefly discuss the architecture of this system. In particular, we address its three major components: the stereo vision sensor, the automated model-fitting algorithm, and the user interaction techniques we have employed.

In addition, we describe the design, execution, and results of experiments that evaluated the performance of our interactive modeling system. The experiments utilized human subjects. These subjects were trained to pilot the application and then evaluated in their use of the interactive model fitting application compared to their use of an identical but purely manual modeling system. Performance was measured in savings of total task time and operator effort as well as accuracy (pose and closeness of fit to the range data).

III. BACKGROUND

In the past, the creation of site models has typically been a primarily manual or automated task. Predominately manual methods rely on a human operator as opposed to a computer for creating and placing object models corresponding to objects observed in a scene. One such method used a gripper attached to robotic arm to seize and place objects for further inspection, analysis, and eventual addition to a model[3]. In Sandia's Graphical Programming System, the location of an object was determined by manually operating a robot to make it touch the object with a probe[4]. These and other manual site and object modeling methods give accurate results, but are often tedious and slow.

Primarily automatic systems are designed to overcome the lack of speed exhibited by manual systems. The computer vision field offers several techniques for automatically recognizing and creating models from range data. We wish to distinguish between two of these techniques: systems that construct a surface map[5] or occupancy map[6] of a scene; and those (like ours) which fit geometric models to discrete objects in the scene. In the latter class of techniques, systems have been developed to fit superquadric volumetric primitives[7], generalized cylinders[8], and parametric geons[9], to range data. Hebert, *et al.* used the operator to select an initial region of interest in the image, then automatically fit a cylinder to the surface data[10].

In other systems, specific object models have been fit to range data. For example, Grimson, *et al.* matched a model derived from MRI data to laser range data[11]. Besl and McKay registered two 3-D shapes using the iterative closest point algorithm[12]. Many techniques use an iterative algorithm to find a solution, either based on least squares or Kalman filtering. One such system used these methods for vehicle docking[13].

Researchers contest that fully automated systems for object recognition and modeling are incapable of matching the human's ability to employ background knowledge, common sense, and reasoning. For instance, automated techniques have difficulty modeling occluded objects correctly. Occlusion is a common trait of unstructured environments such as a drum that is half-buried in a landfill.

Recently, there has been a movement toward interactive applications in robotics. Interactive modeling applications combine automated object recognition techniques with human supervision and incorporation of *a priori* knowledge in an effort to overcome obstacles like occlusions, sparse and noisy data, and segmentation of objects from the background. DARPA's ongoing RADIUS project is an example of how interactive systems have been used successfully to build site models in the domain of overhead image interpretation. In one such system, much work has gone toward automating portions of the interpretation process using image understanding techniques[14]. Interactive techniques have been successfully developed to extract building and road models[15]. For example, an image analyst can provide an initial approximation for a building outline and let the vision system perform local optimization of the variables to fit the data[16]. Although promising, these methods have only been applied to the overhead image interpretation domain (primarily 2-D). They must be modified to apply them to the robotic domain (primarily 3-D) and other environments in which viewpoints are not as constrained.

The following section describes our interactive site modeling system. Our system utilizes automated fitting of object models in conjunction with human supervisory control to create site models quickly and accurately and to overcome the problems presented by occlusion, segmentation, and sparseness of range data.

IV. SYSTEM DESCRIPTION

Our system consists of three major components:

1) A stereo vision sensor as a source of range data points.

2) A model fitting algorithm that uses an optimization routine known as simulated annealing.
3) An interactive (supervisory) technique known as traded control.

A. Stereo Vision Sensor

We have developed a stereo vision sensor for obtaining range data. Stereo vision is a technique that matches points between images from multiple cameras and uses triangulation to compute range. Stereo vision typically produces sparse range data because of the difficulty in matching points correctly. We chose stereo because:

1) It is a more challenging sensor. If our techniques work well with stereo, it is highly probable that they will work with active sensors.
2) Stereo has potential advantages of cost, size, and mechanical reliability over active sensors.

However, our system was designed to use any source of range data. In fact, it has already been tested with data from a structured light sensor.

The stereo vision system detects "interesting" points (often in areas of high contrast) in the left and right images and matches them using a cross-correlation technique. Each candidate match determines a point in 3-D. Figure 1 shows the range points as cross-hairs overlaid on the left image.

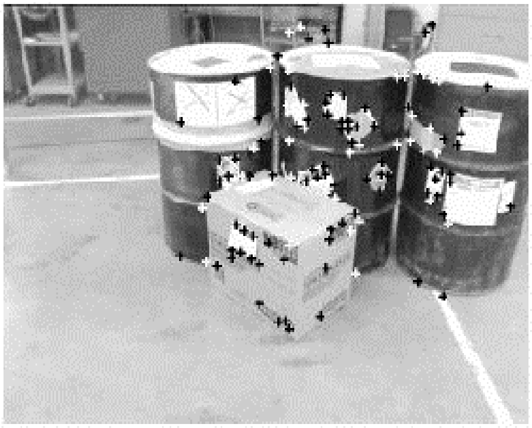

Figure 1 Range data points (crosshairs) from the stereo vision sensor.

We have tested the system on 13 real scenes and numerous synthetic scenes. The result was an average of 88 3-D points per scene. Then, we manually examined these 3-D points and found an average of 4 faulty matches remaining per scene (*i.e.*, points with a large range error).

B. Model Fitting Algorithm

The next step was to develop an algorithm to fit geometric primitive models to range data. The algorithm iteratively adjusts a state vector representing the parameters of the model in order to minimize the error distances between the range data points and the nearest point on the visible surface of the model. Currently, the state vector represents only the location and orientation (6 degrees of freedom, or DOF) of the model, but we are in the process of extending it to include parameters describing size and shape. The algorithm uses the "downhill simplex" method to search for the minimum[17]. With this method, however, there is a strong possibility that the algorithm will become stuck in a local minimum that is far from the true solution.

To avoid this, a simulated annealing algorithm[18] was built on top of the downhill simplex method. Simulated annealing allows the state vector to occasionally move to a point with higher (worse) error and continue its search for the global minimum. The probability of moving to a higher error state is dictated by a "temperature" parameter: a higher temperature means a higher probability of moving to a point of greater error. The algorithm is analogous to the slow cooling of a metal, which allows the atoms to arrange themselves into a crystal that is the minimum energy state (hence the name "simulated annealing"). Although we reduced the temperature at a constant rate, other reduction schedules are possible and are problem dependent[19].

Another important issue was the definition of the error function. One possibility was to define the error as the sum-of-squared distances between the range data points and the nearest point on the surface of the model. Using this approach, points with large distances would dominate the total error score. This means that outliers (points that do not belong to the model) would greatly affect the resulting fit. Therefore, an error function was needed in which outliers do not affect the fit. This error function was defined as

$$errorscore = \sum_{i=1}^{\#pts} -G/r_i$$

where r_i = distance from the i^{th} point to the model and G is a constant. As r_i approaches infinity, the contribution of the point to the error score goes to zero. As r approaches zero, the contribution of the point to the error score would approach negative infinity; however, we limit the minimum score to $-G/D_{MIN}$, where D_{MIN} is the estimated uncertainty in position of our data points (1 cm). This error function has the same form as the potential energy resulting from a gravitational or electrostatic force. In fact, each point may be considered to exert a force proportional to $1/r^2$ on the model, where r is the distance to the model surface. Very close points exert a large force; and very distant points exert a negligible force. Thus, outliers have a negligible effect on the overall fit.

In preliminary experimentation, we have found that the model fitting algorithm converges to a minimum in several thousand iterations, which takes about 10-15 seconds on a Silicon Graphics Indigo2 workstation (this also includes the time required to draw 3-D graphics). In some cases, the true solution or global minimum error was attained; but in all other cases, an incorrect solution or local minimum was located. In the latter cases, user intervention (see next section) was necessary to reach the correct solution.

C. User Interaction Techniques

To incorporate user interaction and visualization into our site modeling system, we have developed a 3-D graphical user interface on the Silicon Graphics workstation. The software displays range data points (colored 3-D cubes) and geometric primitives (spheres, cones, cylinders, and parallelepipeds) as overlays on top of background images from the stereo cameras. These objects can be displayed in normal 2-D mode or in 3-D mode, using stereo viewing glasses. For an input device, the operator can use the usual 2-D mouse, or a 6 degree-of-freedom (6DOF) device.

The user first selects a model from a database of geometric primitives to fit to the data points (see Figure 2). The user can then manipulate the size and shape of the model by entering data into a text box. At this point, the user can manipulate the model using the mouse or 6DOF device.

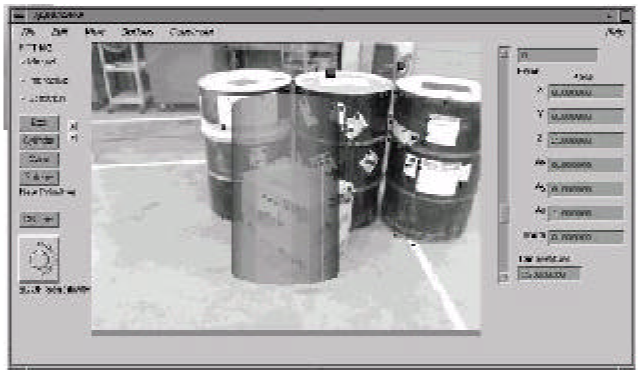

Figure 2 : Initial model position

When the simulated annealing algorithm is running, the model is continuously drawn in the latest estimated position. During the course of the fitting procedure, the model initially jitters around and then jockeys into final position. This animation allows the user to immediately see whether the fitting algorithm is converging to the correct solution and, if not, allows the user to take action to correct the problem. In the meantime, the simulated annealing algorithm automatically reduces the temperature (scaled between 0 and 100) to a fixed percentage of the previous value. After the temperature has fallen to a sufficiently low level, the model stabilizes around its final position. When the temperature reaches 0, the movement stops and the model becomes fixed in its final position (see Figure 3).

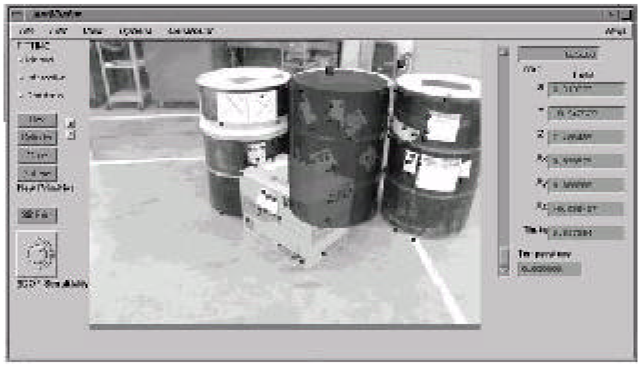

Figure 3 : Final position after fitting

Two metrics are available to the user to gauge the error of the fit of the model to the data points. First of all, the error is continuously displayed as a numerical score. Additionally, the color of the model is changed from shades of red to shades of green as the error reduces.

The technique with which the user can interact with the system to provide guidance is called *traded* control. *Traded* control draws again upon the paradigm of supervisory control[1]. In *traded* control, the user and the system each take turns controlling the pose of the model. This measure has been particularly useful when the model fitting algorithm becomes stuck in a local minimum. Thanks to the continuous rendering of the model, the user can immediately see that this model (shown as a graphical overlay on the image of the scene) has settled into an incorrect pose. By jogging the 6DOF device, the operator can take control of the model from the system and move it towards the correct pose. Releasing the 6DOF device allows the system to resume simulated annealing. We have found that minimal encouragement in the direction of the global minimum was required to push the model out of the local minimum and allow it to automatically find the correct pose. Time spent manipulating the model manually is referred to as "User Time" in the description of experiments that follows.

V. EXPERIMENTS AND RESULTS

We have designed and executed a set of experiments whose purpose was to evaluate the performance of our interactive modeling system compared to a similar, purely manual modeling system. For a fair comparison under experimental conditions, our manual modeling scheme was simply interactive modeling with the automated portion turned off (as was done by Hsieh on the RADIUS project[20]). In our application, manual modeling was accomplished via manipulation of the model with only the 6DOF device. Similar experiments, designed to realize the effect of an operator aid on performance, have been done for a machine-vision-based teleoperation system[21].

A. Hypotheses and Experimental Design

Prior to designing our experimental evaluations, we set forth the following hypotheses concerning the performance of our interactive modeling application:

1) Interactive modeling should outperform purely manual modeling in total task time and fitting accuracy (where fitting accuracy includes both pose error and error in fitting to the range data).
2) Operator effort (measured in terms of user time) decreases significantly when utilizing interactive modeling.
3) Interactive modeling makes up for a lack of expertise. Specifically, subjects who differ greatly in mean task time using purely manual modeling will be statistically indistinguishable on the basis of task time using interactive modeling.
4) Initial placement of the primitive object model by the human supervisor will improve performance when interactively modeling a “difficult” scene, while minimal human interaction (no initial placement) will increase performance on “easy” scenes.

These hypotheses were based on the results of informal evaluations done previously.

The above hypotheses dictated which independent variables would be manipulated in our experiments. The first and second hypotheses required us to split the trials between manual and interactive modeling schemes. Comparison of manual trials to interactive trials would then highlight the difference in performance should our hypotheses hold. As a result, 8 synthetically generated scenes were each fit once interactively and once manually by each subject.

The third hypothesis dictated a division between novice and expert users. We rated expertise solely on the ability to fit object models to range data using purely manual modeling. For an even number of subjects, the division between “expert” status and “novice” was found by simply taking the mean task time for purely manual modeling for each subject and assigning “expert” status to the top 50%.

The fourth and final hypothesis required the introduction of two additional independent variables. One of these was “scene difficulty” which was assessed as the average time it takes to complete the task over all subjects. Again, the toughest 50% would be termed “difficult” while the remainder garnered an “easy” designation. A bimodal distribution of scene difficulty was ensured by varying the number, types, and orientations of objects in a scene as well as the density of range points on and around the objects.

The other variable was the interactive fitting strategy. From the informal evaluations, we were led to believe that initial placement of the graphical model in the region of the object of interest by the human supervisor gave faster task times. Additionally, allowing the automated portion of the interactive modeling application to do most of the work in “easy” scenes also seemed to give speedier task times in those situations. Thus, dividing interactive fitting strategies amongst pairs of similar scenes facilitated a direct comparison of the two for both “easy” and “difficult” scenes. Table 1 summarizes the independent experimental variables.

Table 1 : Independent experimental variables

Variable Name	Possible Values
Modeling Method	manual, interactive
User Expertise	expert, novice
Scene Difficulty	easy, difficult
Fitting Strategy	initial placement, no initial placement

Table 2 : Schedule of scenes encountered during experiments

Scene	**Modeling Method**	**Fitting Strategy**
syn1	manual	
syn3	interactive	initial placement
syn9	manual	
syn10	interactive	no initial placement
syn4	manual	
syn2	interactive	no initial placement
syn11	manual	
syn12	interactive	initial placement
syn2	manual	
syn12	interactive	no initial placement
syn10	manual	
syn1	interactive	initial placement
syn3	manual	
syn9	interactive	initial placement
syn11	manual	
syn4	interactive	no initial placement

B. Execution of Experiments

Several human subjects were necessary for the successful completion of these experimental evaluations. Subjects were brought in on a voluntary basis. In all, 14 subjects, all of which were either undergraduate or graduate students, participated in the experiment. The sampling of students was fairly representative of the School of Mines population: 3 were females; most were right-handed (although the 6DOF device favors neither); all were at least fairly computer literate, most exceptionally so; all had normal (sometimes corrected) vision and depth perception; finally, most were in their early twenties, although some were in their late twenties or older.

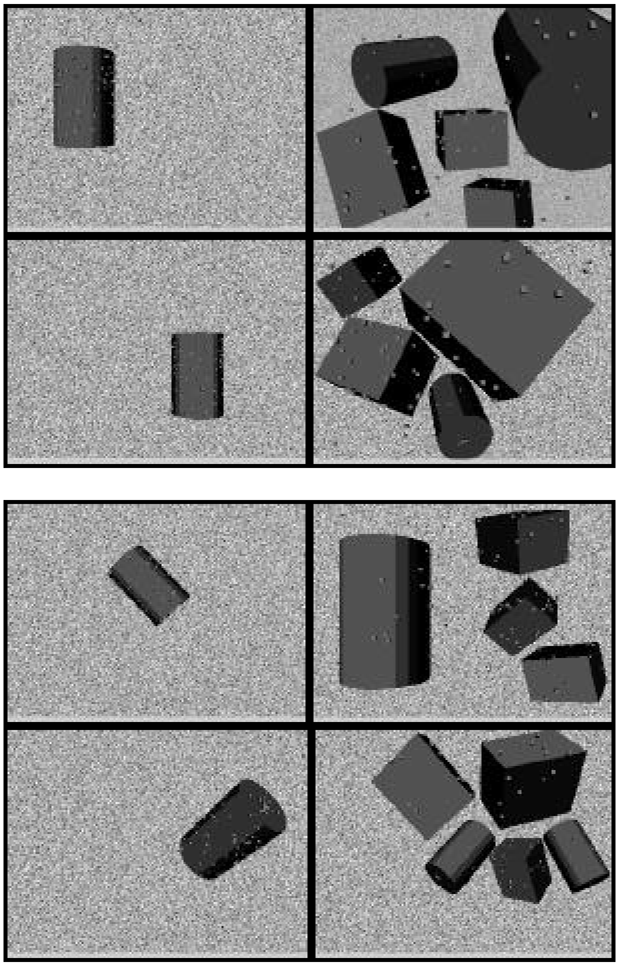

Figure 4 : Scenes (left to right from top) syn1, syn11, syn2, and syn10 (easy); syn3, syn9, syn4, and syn12 (difficult).

Each subject was briefed on the scientific and industrial significance of the system and then was trained for a period of 10 to 15 minutes in the use of the 6DOF device: a Logitech Magellan. The training session covered the three translations (x, y, z), changes of orientation (roll, pitch, and yaw) and allowed the subjects to practice fitting models to separate training data using manual modeling, interactive modeling with initial placement, and interactive modeling without initial placement.

Each subject was then asked to complete the 16 trials. The trials took anywhere from 30 to 50 minutes giving approximately 55 minutes to an hour on average to complete an experiment with a particular subject. The scenes were encountered in a somewhat random order, but were arranged so that alternating trials using manual and then interactive modeling occurred (see Table 2).

Completion of a trial occurred when the user and evaluator agreed that the fit to the data was the closest that could be found. The computer then automatically logged the pose data and the overall and user task times. User time during interactive trials was recorded by the computer every time the 6DOF device was touched and continued until control was returned to the automated modeling routine.

Several factors were kept constant during the experiments in an effort to keep from skewing our results. First of all, all subjects used the same workstation with an equal number of applications running simultaneously. In addition, room lighting was made approximately the same.

The traits of the interactive modeling routine, however, necessitated an additional consideration. In particular, the annealing temperature needed to be lowered to zero before completing the trial to ensure that the model would settle into a minimum error state. This also ensured that every subject had an equal opportunity to fit the data as closely as possible.

C. Results

To determine the validity of our results relative to our original hypotheses, we utilized a statistical comparison called analysis of variance (ANOVA)[22]. Simply put, the purpose of ANOVA is to calculate the confidence with which a particular independent experimental variable is said to affect a dependent variable or desired result. The confidence is found by splitting the experimental trials according to the value of one independent experimental variable and then doing a statistical comparison. For example, the dependent variable "User Time" was shown to have been directly affected by the choice of "Modeling Method" (either manual or interactive). Table 3 summarizes the results of ANOVA applied to data gathered from 224 total trials that were grouped according to the specified independent experimental variables. The results were tested to the 99% confidence level.

Those results which passed the ANOVA validation are summarized in Figure 5 and Tables 4 and 5. From this data, we confirmed the following hypotheses:

1) Interactive modeling should outperform purely manual modeling in total task time and fitting accuracy.
2) Operator effort decreases significantly when utilizing interactive modeling.
3) User expertise does not have a significant effect on interactive modeling task time.
4) Minimal human interaction will increase performance on "easy" scenes.

Table 3 : Summary of ANOVA results

Dependent Variable	Independent Variable	Independent Affects Dependent?
Total Task Time	Modeling Method	Yes
User Time	Modeling Method	Yes
Translation Error	Modeling Method	No
Orientation Error	Modeling Method	Yes
Error Score	Modeling Method	Yes
Task Time (easy interactive scenes)	User Expertise	No
Task Time (difficult int. scenes)	User Expertise	No
Task Time (easy interactive scenes)	Fitting Strategy	Yes
Task Time (difficult int. scenes)	Fitting Strategy	No

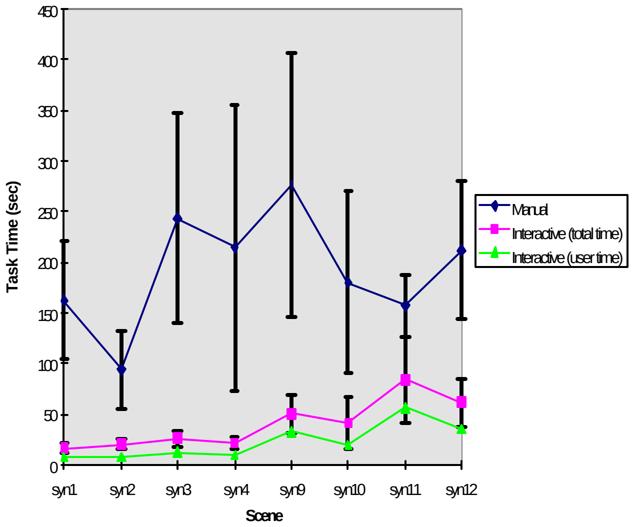

Figure 5 : Comparison of mean task times for each scene.

Also, note that the difference between expert and novice performance for trials where interactive modeling was used does not appear in the results. As shown in the ANOVA table, the confidence that the independent variable "User Expertise" affected the total task time was very low. Therefore, Hypothesis 3 (that "User Expertise" does not significantly affect task time when using interactive modeling) is supported as well. The effect of interactive versus manual fitting on translation error and the performance of initial placement versus no initial placement for difficult scenes failed this criterion, too. This suggests that these portions of the original hypotheses are neither proven nor disproven.

Table 4 : Mean experimental results for all scenes

	Manual	Interactive
Task time (sec)	192.75	40.68
Scaled error score (min 1000.0)	1101.48	1000.62
Orientation error (deg)	1.886	.758

Table 5 : Interactive task times for "easy" scenes using various fitting strategies.

Fitting Strategy	Mean Task Time (seconds)
Initial placement	50.75
No init. placement	31.14

VI. CONCLUSIONS AND FUTURE WORK

We have developed an interactive system for creating site models using range data. This system consists of three major components: a stereo vision sensor, an automated iterative optimization algorithm known as simulated annealing, and a supervisory control measure allowing for human interaction. Through formal experimentation and evaluation, we have shown this interactive system outperforms an identical but primarily manually driven application in speed, accuracy, and reliability. Specifically, the following hypotheses were confirmed:

1) Interactive modeling should outperform purely manual modeling in total task time and fitting accuracy.
2) Operator effort decreases significantly when utilizing interactive modeling.
3) User expertise does not have a significant effect on interactive modeling task time.
4) Minimal human interaction will increase performance on "easy" scenes.

We have also demonstrated, in general, that such a system can provide a high-performance solution to the crucial site-modeling problem facing engineers whose robots work in unstructured environments.

In the future, we would like to integrate this system within a robotic test environment so that we can work to develop efficient representations of the site model that are easily understood by robotic systems. In addition, we intend to add scaling and other "morphing" attributes to our degrees of freedom to increase the flexibility of our application by offering a wider variety of available primitives to fit to the scenes.

VII. ACKNOWLEDGMENTS

The authors wish to acknowledge the contributions of Torsten Lyon, Khoi Nguyen, Rex Rideout, Doug Swartzendruber, and all of the volunteer test subjects. We also wish to thank Mark McKay of INEL for his guidance and assistance in this project. Finally, we wish to thank Charles Little of Sandia National Labs for providing us with structured light range data.

REFERENCES

[1] T. B. Sheridan, *Telerobotics, automation, and human supervisory control*, Cambridge, Massachusetts, MIT Press, 1992.

[2] S. Lee, "Intelligent Sensing and Control for Advanced Teleoperation," *IEEE Control System Magazine*, Vol. 13, No. 3, pp. 19-28, 1993.

[3] D. Cannon, G. Thomas, C. Wang, and T. Kesavadas, "A virtual reality based point-and-direct robotic system with instrumented glove," *Int'l J. of Industrial Engineers - Applications and Practice*, Vol. 1, No. 2, pp. 139-148, .

[4] M. J. McDonald and R. D. Palmquist, "Graphical programming: On-line robot simulation for telerobotic control," *Proc. of International Robots and Vision Automation Conference*, pp. 22-59..22-73, 1993.

[5] A. Leonardis, A. Gupta, and R. Bajcsy, "Segmentation of Range Images as the Search for Geometric Parametric Models," *International Journal of Computer Vision*, Vol. 14, No. 3, pp. 253-277, 1995.

[6] Y. Roth-Tabak and R. Jain, "Building an Environment Model Using Depth Information," *Computer Magazine (IEEE)*, Vol. 22, No. 6, pp. 85-90, 1989.

[7] R. Bajcsy and F. Solina, "Three dimensional object representation revisited," *Proc. of First Int'l Conf on Computer Vision*, IEEE Computer Society, London, pp. 231-240, 1987.

[8] T. O'Donnell, T. Boult, X. Fang, and A. Gupta, "The extruded generalized cylinder: A deformable model for object recovery," *Proc. of IEEE Conf. Computer Vision and Pattern Recognition*, pp. 174-181, 1994.

[9] K. Wu and M. Levine, "Recovering parametric geons from multiview range data," *Proc. of IEEE Conf. Computer Vision and Pattern Recognition*, pp. 159-166, 1994.

[10] M. Hebert, R. Hoffman, A. Johnson, and J. Osborn, "Sensor-Based Interior Modeling," *Proc. of Robotics and Remote Systems*, American Nuclear Society, Monterey, CA, pp. 731-737, 1995.

[11] W. E. L. Grimson, et al, "An automatic registration method for frameless stereotaxy, image guided surgery, and enhanced reality visualization," *Proc. of*

Computer Vision and Pattern Recognition, IEEE, pp. 430-436, 1994.

[12] P. J. Besl and N. D. McKay, "A Method for Registration of 3-D Shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 14, No. 2, pp. 239-256, 1992.

[13] E. D. Dickmanns and V. Graefe, "Applications of Dynamic Monocular Machine Vision," *Machine Vision and Applications*, Vol. 1, No. pp. 241-261, 1988.

[14] S. J. Gee and A. M. Newman, "RADIUS: Automating Image Analysis Through Model-Supported Exploitation," *Proc. of DARPA Image Understanding Workshop*, Morgan Kaufmann, pp. 185-196, 1993.

[15] S. Heuel and R. Nevatia, "Including Interaction in an Automated Modeling System," *Proc. of Image Understanding Workshop*, ARPA, Palm Springs, pp. 429-434, 1996.

[16] P. V. Fua and Y. G. Leclerc, "Model driven edge detection," *Proc. of DARPA Image Understanding Workshop*, Morgan Kaufmann, pp. 1016-1021, 1988.

[17] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C*, 2nd ed., Cambridge University Press, 1992.

[18] S. Kirkpatrick, J. C.D. Gelatt, and M. P. Vecchi, "Optimization by Simulated Annealing," *Science*, Vol. 220, No. 4598, pp. 671-680, 1983.

[19] K. H. Hoffman and P. Salamon, "The Optimal Simulated Annealing Schedule for a Simple Model," *Journal of Physics A: Mathematical and General*, Vol. 23, No. 15, pp. 3511-3523, 1990.

[20] Y. Hsieh, "Design and Evaluation of a Semi-Automated Site Modeling System," *Proc. of Image Understanding Workshop*, Vol. 1, Morgan Kaufmann, Palm Springs, CA, pp. 435-459, 1996.

[21] W. Hoff, L. Gatrell, and J. Spofford, "Machine Vision Based Teleoperation Aid," *Telematics and Informatics*, Vol. 8, No. 4, pp. 403-423, 1991.

[22] R. R. Johnson, *Elementary Statistics*, 3 ed., Boston, Duxbury Press, 1980.