VECIMS 2003 - International Symposium on
Virtual Environments, Human-Computer Interfaces, and Measurement Systems
Lugano, Switzerland, 27-29 July 2003

Toward an Occupancy Model Rendering Framework for Robotic Systems Operating under Human Supervision

P. Payeur
Vision, Imaging, Video and Audio Research Laboratory
School of Information Technology and Engineering
University of Ottawa
Ottawa, Ontario, Canada, K1N 6N5
ppayeur@site.uottawa.ca

Abstract* - The introduction of 3-D modeling with probabilistic encoding of occupancy state of space leaded to the development of powerful tools to generate rich representations of environments. But the sophistication of the modeling scheme introduces important concerns in the way such representations can be efficiently displayed to human operators for validation purpose and decision making in the guidance of robotic systems. This paper proposes a framework and describes the characteristics of a prototype that has been designed for probabilistic occupancy grids rendering with the objective to alleviate the relative complexity associated with the visual interpretation of 3-D probabilistic occupancy grids. Various aspects of the interaction with the operator are investigated and suitable approaches are identified to provide maximum flexibility. Examples are presented of an application of the proposed rendering framework in telemanipulation operations.

I. INTRODUCTION

The development of integrated robotic systems able to operate in unknown environments implies important processing efforts on data collected by means of various sensors. 3-D representations of the occupancy state of space for complex environments must also be built. Powerful tools have been developed to achieve this tasks using the paradigm of probabilistic occupancy grids [2, 5, 7]. But in spite of significant research efforts invested over the last decade, registration and data fusion still remain critical issues that need to be refined. As a result, robotics systems able to operate in a completely autonomous way when dealing with unknown or dynamic environments are not yet achievable except for very specific and simple applications.

However, promising semi-autonomous robotics systems can already be developed provided that human supervision is ensured in the overall process. In such systems, computers and artificial intelligence are used to handle heavy computational tasks such as data processing, path planning and collision detection while human operators ensure the validation of intermediate results after each step of operation and make high level decisions. A nice example of such a semi-autonomous telerobotic system is currently under development for safe maintenance of electric distribution networks in Canada as illustrated in figure 1 [1]. In this application, robots are used for insulators replacement on live electricity networks.

As each electrical pole configuration differs, a model of the scene must first be built from data fusion between measurements collected with a range finder from different viewpoints. This modeling phase must be completed before allowing the robots to interact with the objects in the environment while avoiding undesired collisions with high-voltage components. In spite of the level of automatization introduced in the system, human operator keeps an important role in the procedure as he has to estimate the reliability and the completeness of the model to validate the virtual representation before it can be used by the path planning module. Actually, the human operator keeps a decisional authority on all internal modules as shown in figure 2.

This necessary interaction between the human operator and the modeling tool requires that an efficient and intuitive display of the environment representation is provided to the operator. Since the model is progressively refined as sensors are moved around the area of interest, the rapidity, the quality and the flexibility of the representation are critical to allow the operator to validate the completeness and the correctness of the model. He might also eventually have to directly interact with the model by selecting some areas where the model must be refined and by specifying points or objects with which the robotic devices must interact (high level recognition).

In such a context and under the probabilistic modeling paradigm, strategies need to be identified to provide efficient rendering of 3-D occupancy models in an intuitive way that human beings with a minimum of training can quickly interpret. It is important to notice that, in this context, the model is mainly dedicated to be interpreted by artificial intelligence algorithms. As a consequence, the goal is to built and display representations that are rich in terms of their content rather than nice looking virtual models as found in virtual or augmented reality applications.

This paper describes a rendering framework developed to overcome the relative complexity associated with 3-D probabilistic occupancy grids interpretation. The main factors that have been considered are: the elimination of ambiguities in the representation in order to facilitate and accelerate the validation process, the update rate of the representation, and the possibility for the operator to interact with the model to specify the parameters of the next task to perform.

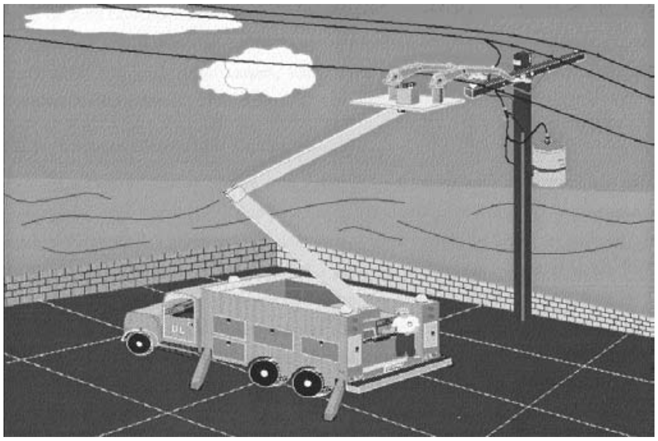


Fig. 1. Semi-Autonomous Robotic System for Live Electricity Line Maintenance.

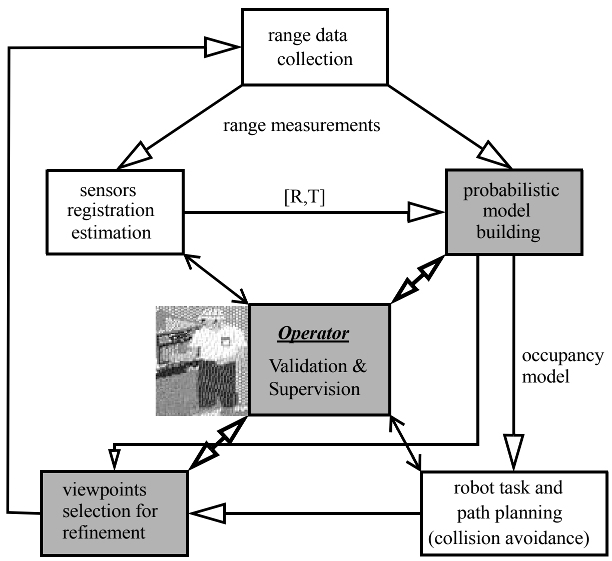


Fig. 2. Structure of the robotic system operating under human supervision.

The following sections summarize the main challenges that have been faced in the development of the human interface and discuss the solutions that have been implemented. This prototype is part of a complete modeling and robot control tool currently under development.

II. DESCRIPTION OF MODELS

The type of models that have been selected for the development of the present telerobotic system is that of 3-D probabilistic occupancy grids encoded under the form of octrees. This choice has been made on the basis of the nature of the information that these models encode, that is the cluttering state of space rather than surface or texture definitions. As models are dedicated to robot guidance among obstacles, a volumetric representation appears to be more suitable.

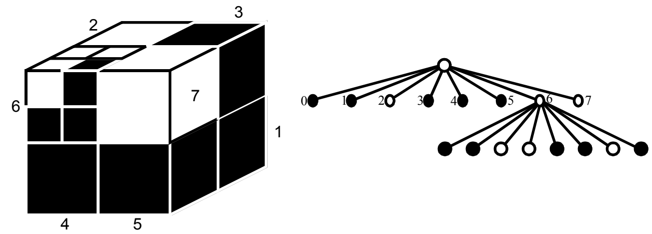


Fig. 3. Multiresolution occupancy grid and equivalent octree.

Probabilistic occupancy grids are built from the tessellation of 3-D space into voxels by means of a recursive subdivision. Each voxel is tagged with the probability that the corresponding area is occupied (between 0% and 100%) [2, 10]. Octrees are a perfectly equivalent volumetric representation as they correspond to an identical subdivision of space but store the information in a hierarchical structure that offers significant advantages for accessing data. Even though the explicit information about geometrical relationships between voxels is lost in octrees, it can be easily retrieved by a simple tree traversal or an appropriate numbering scheme used with specific algebra [8, 11].

Multiresolution occupancy grids are a variation of standard occupancy grids. Multiresolution grids contain voxels of various sizes, or levels of resolution, encoded in the same model. This property significantly contributes to make probabilistic modeling schemes tractable for 3-D space has large volumes with identical occupancy state can be merged in a single voxel and processed in a single step. Data fusion can therefore be accelerated for environments with uniform distributions. Figure 3 shows the relationship between a simple multiresolution occupancy grid in 3-D space and the corresponding octree encoding.

III. DESCRIPTION OF THE PROPOSED RENDERING FRAMEWORK

The rendering framework that has been developed in parallel with the 3-D probabilistic modeling scheme, relies on some resources from the OpenGL graphical programming language [6] and the XForms graphical user interface development library [12]. Our investigation with the model rendering tool revealed a given number of aspects that must be considered to provide human operators with an efficient interface when dealing with tridimensional multiresolution probabilistic representations. It is mandatory that the interface succeeds to display meaningful information that supports the operator in the process of validation and decision-making independently of the complexity of the environment. This section analyzes the various aspects that have been investigated.

A. Viewpoint selection

As 3-D environments are to be represented, the operator must have the possibility to select a desired viewpoint to observe the scene from a proper angle that allows him to assess the quality of the model, the boundaries of objects or the availability of sufficient empty space for the robot to navigate in a given area. This function is currently achieved by means of a series of cursors that are manipulated one at a time. Various schemes have been tested such as controlling a set of three translations and three rotations, as well as moving the viewer or moving the object. We concluded that moving the object resulted in a more comfortable interaction and that the number of parameters to adjust was advantageously minimized by a set of three translations (left-right, up-down, back-forth) and only two orientations (azimuth and elevation). For orientation selection, the representation of composed transformations based on three rotations around fixed axes is quite difficult to manage for human beings while the manipulation of azimuth and elevation is much more intuitive. It did not appear necessary to include a swing parameter as it does not improve the viewing and the estimation of the 3-D scene in general and rather tends to confuse the operator about his current spatial location with respect to the object.

B. Displaying of unknown areas

Working with probabilistic occupancy grids rather than deterministic ones that have only three possible occupancy states (*empty*, *occupied* or *unknown*) brings important issues in the way the rendering of voxels is managed. Especially, the fact that the telerobotic system is designed to work in undefined environments implies that the modeling system must built a representation of space cluttering without any *a priori* knowledge on the scene. The initial probability level of the model for the selected volume of interest (typically 1 m^3 in this application) is then set to 50% everywhere, which corresponds to an unknown state of space. As the sensors collect information about the configuration of objects in 3-D space and the modeler merges these data, the unknown area is progressively reduced and replaced by the occupancy state of that part of the environment that is being modeled.

However, it is in general impossible to simultaneously cover all the space with limited field of view sensors. Moreover, unknown cells are also created in the model when contradictory data are merged and lead to an undetermined state of space for some voxels. As a result, unknown cells remain in the model and the operator must face situations where unknown areas are dominant in the model. The main consequence of these unknown cells is that, if they are included in the rendering, they place relevant parts of the scene in occlusion from several viewpoints that the viewer might select as shown in figure 4. Interpretation of the model might then be dramatically degraded.

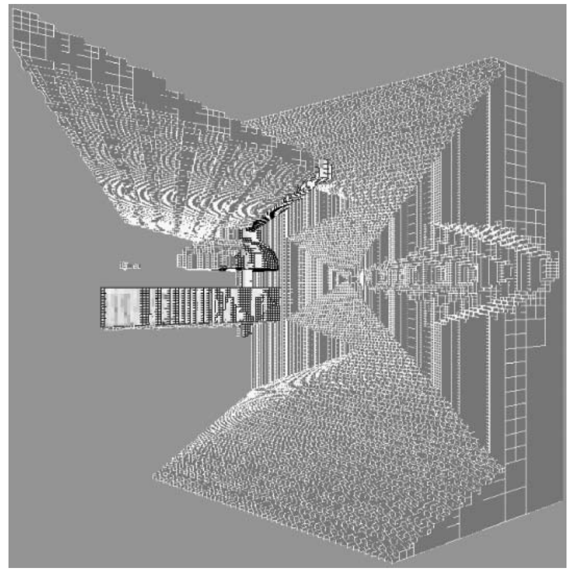


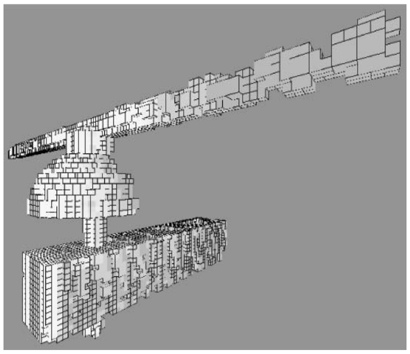
Fig. 4. Occupancy model rendering with unknown cells.

In order to avoid such situations, a function for unknown cells removal is integrated in the rendering framework. It is proposed as an option that the operator might activate to obtain clear views of the part of the environment for which measurements have been collected and integrated in the representation. In practice, this function should be activated for most part of the operation period. It ensures that all voxels having an occupancy probability around 50% are eliminated from the display and the corresponding volume is left empty.

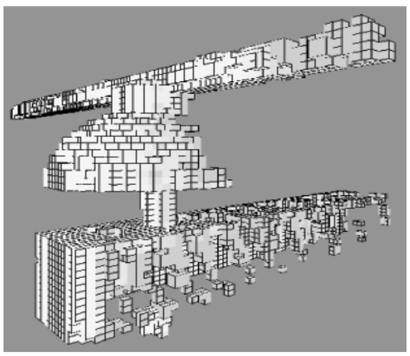
C. Minimum level of occupancy probability

Experimentation also revealed that removing unknown cells is not sufficient to ensure clear observation of the current state of the probabilistic model. Voxels tagged with a low probability of occupancy (associated to empty space) also tend to clutter the display and preempt the visualization of real objects that are rather associated with high probability of occupancy. Therefore the operator must have the possibility to select a minimum level of occupancy probability above which voxels are to be displayed.

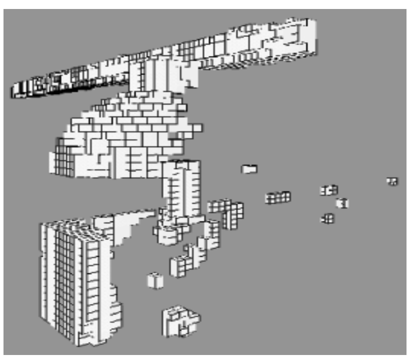
This functionality provides another advantage. By adjusting the threshold of minimum occupancy probability to be displayed, the operator has the possibility to isolate some cells having a probability above the selected threshold. This allows verification on the shape of objects up to the maximum level of resolution for which the model has been computed. Figure 5 shows a same model of the extremity of an electrical pole under various minimum levels of occupancy, respectively 65%, 80% and 95%. For clarity purpose, gray levels are associated with different occupancy probabilities (black being 0%, white being 100%).



$P(OCC)_{min}$ = 65%

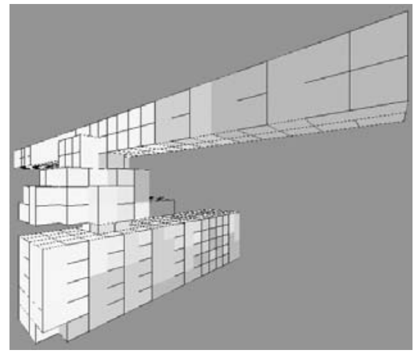


$P(OCC)_{min}$ = 80%

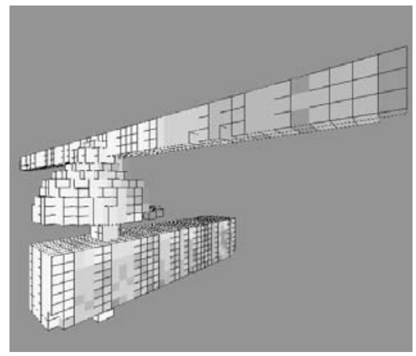


$P(OCC)_{min}$ = 95%

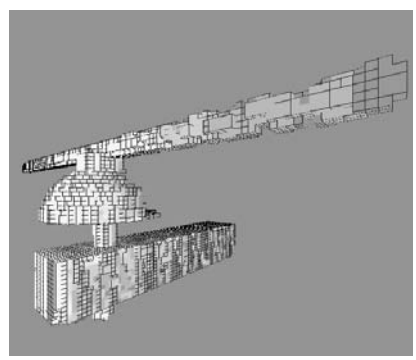
Fig. 5. Occupancy model rendering for various minimum levels of occupancy.



resolution = 31.3 mm



resolution = 15.6 mm



resolution = 7.81mm

Fig. 6. Occupancy model rendering for various resolutions.

We observe that the number of cells that are actually displayed is reduced as the minimum level of occupancy is increased. This phenomena is explained by the fact that some regions have been scanned only once by the sensor while others have been measured many times due to the overlap between the fields of view of sensors. As the occupancy probability exhibits a progressive validation characteristic (points measured many times get more confidence than those measured only once), the occupancy probability tends to increase as a function of the number of measurements collected on the surface of objects [7]. The reverse effect is also observed for empty space as the occupancy probability is decreasing as a function of the number of measurements that confirms that a given point is free of any object.

D. Level of resolution for validation

Another aspect that needs to be considered to provide the operator with a functional interface is the maximum resolution of display that he requires to validate the model and to make decisions. The goal of the present system being not to generate nice looking virtual representations of objects, a maximum resolution does not appear to be required under all circumstances. Hence, the resolution of the displayed model influences the refresh rate. As a consequence, the interface can easily become unpleasant to use if high resolution is to be displayed by default. The operator should then have the possibility to dynamically select the maximum resolution at which he wishes to examine the model. This function reveals to be useful when changing the viewpoint as the operator might reduce the resolution to quickly determine the best virtual viewpoint for the task he wants to perform and then increase the resolution to retrieve a more detailed representation. This selection is controlled by a slider that steps through all possible levels of resolution available in the model, those being defined as powers of 2 since the grid results from a recursive subdivision of space along each axis. Figure 6 shows a same model of the insulator and beam assembly at various resolution levels.

E. Transparency versus external observer

Intuitively, the representation of a 3-D occupancy model dedicated for robot path planning and collision avoidance should follow the idea of space traversal. That is, if the robot is to be guided through obstacles based on this virtual representation of space, the operator should be allowed to make a similar travel in the virtual world. Toward this goal, experiments have been conducted with transparency effects in the rendering of the occupancy grid. The idea is to represent empty voxels as transparent regions while occupied voxels (with a high level of confidence) are displayed as opaque areas. Voxels with intermediate occupancy values see their transparency level adjusted accordingly as shown in figure 7. This approach offers the advantage that the full content of the model can be presented to the operator while having him control a minimum set of options. It might also significantly help the operator to guide a robot through cluttered space in real-time. However, experimentations revealed that the interpretation by human beings of such a representation is extremely demanding and difficult, especially because of the fuzziness of the drawing that hardly allows to precisely determine the boundaries of objects. As a consequence, operators tend to get tired after a relatively short period of interaction with the interface. This functionality still remains available in the current version of the prototype but is not widely used.

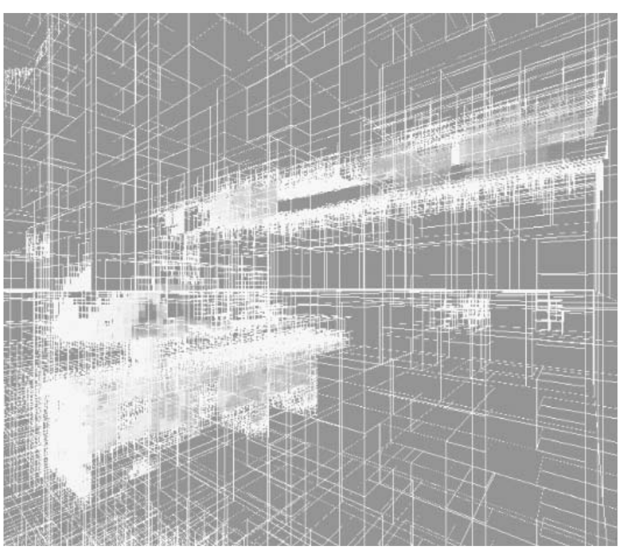


Fig. 7. Occupancy model rendering with transparency.

F. Rendering techniques

Displaying a virtual scene described with elementary polygons is different than rendering the corresponding 3-D probabilistic occupancy model as illustrated in figure 8 where a polyhedral model of the extremity of an electrical pole is represented along with the corresponding probabilistic occupancy model. In the occupancy representation, the number of elementary components to draw (cubes of the occupancy grid) is significantly larger than the number of primitive geometrical shapes (cylinders, spheres and parallelepipeds) in the polyhedral representation. As a consequence, problems of occlusions associated with the selected viewpoint and the speed of rendering significantly differ.

However, the update rate of the model remains one of the main concerns as the occupancy representation needs to be built from scratch each time the telerobotic system is used for an intervention. The human operator can initially drive the sensors toward the area to be modeled in manual mode but from that point he requires some feedback about the current state of the representation in order to help it improve its completeness and reliability. A progressive refinement scheme is then used to guide the sensory devices around the objects with which the manipulator will have to interact once a global representation of the scene is achieved.

As the prototype of the system has been developed on the basis of OpenGL graphical language, Z-buffering has been used to draw the volumetric representation [6]. Basically, all voxels at maximum resolution (those that do not have children) included in the model were drawn and the interface relied on the Z-buffering functionality of OpenGL to plot only the visible polygons on the display. This strategy which appears to be sufficient in 2-D or with very simplistic 3-D models quickly revealed its limitations for complex environments represented at high resolution.

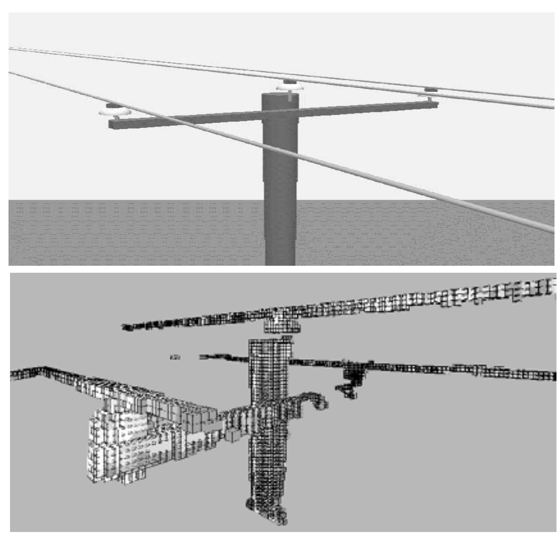


Fig. 8. Virtual representations of a scene with polyhedral structure and the corresponding probabilistic occupancy grid.

Alternatives such as ray-tracing [3] have been considered to improve performances of the rendering tool. But the computational workload required to estimate projections from the viewpoint and to identify voxels that need to be displayed does not seemed justified, especially because of the nature of the representation that can easily contain more than one million voxels at various resolutions since relatively large areas of space need to be modeled.

An approach of prioritized-layered projection that relies on an occupancy-based tessellation of space and that is especially well suited for high depth complexity models has been investigated [4]. This strategy is efficient for time-critical rendering applications but is not conservative and tends to distort the reality. Given the inherent complexity already associated with the interpretation by the human operator of the correct representation of the scene, it is not suitable to introduce supplementary distortion due to the rendering scheme.

Most algorithms proposed for octree rendering also assume that the tree has a subset of polygons at every leaf. In the present case, only a cube is present at every leaf. Therefore, the approach can be simplified based on this particularity. Moreover, taking advantage of the recursive structure of the multiresolution octree that intrinsically encodes the geometrical relationships between cells provides another interesting optimization opportunity [9]. Using a fast neighbor finding technique that has been developed for this type of octree encoding [8] and that allows to identify neighbor cells in spite of their different sizes, a direct scanning of the octree structure can be performed in order to determine if some sides of a given cube are visible from outside the model. Proceeding to a single traversal of the octree and verifying if each of the 6 faces of the cube has a neighbor or not (a cell tagged with a probability above the selected minimum level of occupancy) allows to create a limited list of polygons (squares) that might need to be displayed. This approach significantly reduces the number of polygons to draw as cubes lying inside of the model are automatically removed while only a subset of the faces of those cubes that are visible is kept. Combining this approach with OpenGL Z-buffering functionalities significantly speeds up the rendering process and makes the interaction with the interface more interesting.

The characteristics previously described have been integrated in a generic graphical tool that has been used over the last few years for the development of the 3-D probabilistic modeling scheme. Other intermediate results can also be displayed by the same tool but are not discussed in the present paper as they do not represent such a critical aspect for the interaction with human operators during robotic interventions.

IV. CONCLUSION

A rendering framework adapted to multiresolution probabilistic occupancy models for 3-D space and dedicated to serve as a dynamic interface with human operators for telerobotic applications has been described. Conclusions from an investigation of various strategies to efficiently display complex models to the operator for validation and decision making is presented. Approaches that revealed to be the most suitable have been identified and implemented in a prototype. As a result, this work provides an insight into an important aspect of probabilistic modeling which is often neglected in the development of modeling strategies. It demonstrates that probabilistic representation of space occupancy can be efficiently interfaced with human beings provided that some supplementary functionalities and flexibility in the display parameters are provided to the operator. As this prototype is still under development, further adjustments are expected to bring the interaction more friendly and comfortable, especially by improving the interaction for viewpoint selection and by increasing the rendering speed.

REFERENCES

[1] M. Boyer, “Systems Integration in Telerobotics: Case Study: Maintenance of Electric Power Lines”, in *Proc. of the IEEE International Conference on Robotics and Automation*, vol. 2, pp. 1042-1047, Minneapolis, MN, April 1996.

[2] A. Elfes, “Using Occupancy Grids for Mobile Robot Perception and Navigation”, *IEEE Computer*, 22(6), pp. 46-57, 1989.

[3] J.P. Jones, “Real-Time Construction of Three-Dimensional Occupancy Maps”, in *Proc. of the IEEE Int. Conf. on Robotics and Automation*, vol. 1, pp. 52-57, May 1993.

[4] J.T. Klosowski, C.T. Silva, “The Prioritized-Layered projection Algorithm for Visible Set Estimation”, in *IEEE Transactions on Visualization and Computer Graphics*, pp. 1077-2626, Apr.-Jun. 2000.

[5] H. P. Moravec, *Robot Spatial Perception by Stereoscopic Vision and 3D Evidence Grids*, Technical Report CMU-RI-TR-96-34, Robotics Institute, Carnegie Mellon University, Pittsburgh, Pennsylvania, 1996.

[6] J. Neider, T. Davis, M. Woo, *OpenGL Programming Guide, Addison-Wesley*, Reading, MS, 1993.

[7] P. Payeur, D. Laurendeau, C. M. Gosselin, “Range Data Merging for Probabilistic Octree Modeling of 3-D Perturbed Workspaces”, in *Proc. of the IEEE Int. Conf. on Robotics and Automation*, vol. 4, pp. 3071-3078, Leuven, Belgium, 1998.

[8] P. Payeur, “Efficient Neighbor Identification in Multiresolution kD-Trees for Mobile Robot Navigation”, in *Proc. of the IASTED Int. Conf. on Robotics and Automation*, Clearwater, FL, Nov. 2001.

[9] I. Provaznik, D. Schwarz, P. Fedra, “Complex-Valued Multiresolution Volume Rendering in 3D Imaging”, *Computers in Cardiology*, pp. 433-435, 2001.

[10] H. Samet, *The Design and Analysis of Spatial Data Structures*, Addison-Wesley, Reading, MA, 1990.

[11] H. Samet, “Neighbor Finding in Images Represented by Octrees”, *Computer Vision Graphics Image Processing*, vol. 46, pp 367-386, 1989.

[12] T.C. Zhao, M. Overmars, *Forms Library, a Graphical User Interface Toolkit for X*, 1996.