MVA'90 IAPR Workshop on Machine Vision Applications Nov. 28-30,1990, Tokyo

USING AND GENERATING ENVIRONMENT MODELS FOR INDOOR MOBILE ROBOTS

Yuval Roth-Tabak and Terry E. Weymouth
The Artificial Intelligence Laboratory
The University of Michigan

1101 Beal Ave.
Ann Arbor, MI 48109-2110, USA

ABSTRACT

An autonomous mobile robot must be able to combine uncertain sensory information with prior knowledge of the world. Moreover, these operations have to be performed fast enough for the robot to be able to react to the changes in the world. This paper presents a model-driven system for a real-time indoor mobile robot. As the robot is constantly in motion, information from an Environment Model is used to anticipate information-rich features and to direct selective sensing. This model-based expectation helps the system overcome problems of slow sensing rate by requesting only that information which is immediately necessary. Different types of features activate different sensing modes. Uncertain sensor information is combined with prior World Model knowledge to reduce uncertainty in the model. We present a hall-following robot, based on these principles, which exhibits real-time navigation performance. It does this despite primitive and relatively slow sensing. As such a system is dependent on an existing model, we have also examined ways to create similar models from the sensory information. We experimented with a neural network which is trained by back-propagation to identify the desired features in the corridor. A simulation of this system is run off-line, and preliminary results are presented in which a model of a corridor is directly derived from sensory information.

INTRODUCTION

An autonomous mobile robot must be able to combine uncertain sensory information with prior knowledge of the world. Moreover, these operations have to be performed fast enough for the robot to be able to react to the changes in the world. A robot should use a-priori model information to control its sensing. This can be viewed as a conformation process in which the system uses the available model information to minimize the perceptual processes, and to extract only that information which is needed to verify the existing structures in the model. The use of the model will help the system avoid unnecessary operations; however, a totally 'model-trusting' system may exhibit hallucinatory behavior when a mismatch evolves between the reality and the model. The system will then try to force instantiations of model structures even when the structures do not exist in reality. To avoid these problems, a system must also maintain a constant check on the disparity between expected features and actual sensed features. If the disparity is too high, the system should adapt by modifying its perception and action strategies, possibly slowing down, and building a model directly from sensory information. This strategy will require more exhaustive information collection and analysis, and, thus, will proceed slower. However, once the information is collected for a particular environment, the system can switch back to the faster model-based strategy.

Earlier robot systems such as[1,2] approached the problem of sensory guided navigation by moving the robot very slowly, actually stopping it altogether so it had enough time to analyse all the information, to arrive at the correct conclusion, plan its path and then only to proceed. This is obviously not a desirable solution if the robot is to perform any task in a dynamic environment since the temporal sampling rate is too low.

At the other end of the spectrum, systems capable of real-time reaction[3], are designed to accommodate this demand by reducing the analysis of information to the most simple operations performed only on windows from the original data set. The problem is that these systems require a precise dynamic model of the environment, and of the statistical characteristics of corruption in the data. A Kalman filter is used to combine new observations with the state of the system. The Kalman filter is an iterative least squares approach and thus will not be able to handle totally erroneous data. A Kalman filtering approach is also used in[4], here, analysis of stereo images is used to guide a robot in a corridor. A real-time obstacle avoidance scheme using sonar data is presented in[5]. Sonar data from successive readings is combined to create a 'histogram grid' with 'obstacle densities' which is in tern used to successfully control a mobile platform. Nevertheless, their system does not incorporate any a priori information, and does not perform any model-based analysis.

Another approach[6] introduces reactive systems with no apparent cognitive parts. The attractive aspect of these systems is the 'subsumption architecture' which provides a framework for placing more complex behaviors on top of more primitive ones. However, these systems lack information analysis capabilities, and behavior is fully hard-wired. The system cannot deal with any 'unexpected' situations, nor has it any means for filtering out inconsistent data.

A general-purpose architecture for controlling mobile robots, the Task Control Architecture, is presented in[7]. This work concentrates on the definition of the control mechanisms, but fails to introduce a distinction between acquired knowledge and available knowledge (see also[8]).

We follow closely the paradigm presented in [9] in which the Environment Model is placed at the heart of a system. This should allow assimilation of information from various sources and over time.

A representation hierarchy for robust navigation and mapping for large scale terrain is presented in[10]. In our system we implement the Sensorimotor, the Topological and the Metric levels discussed there.

We present a system which combines uncertain sensory information, a priori world model information, and expectations from the previous environment model state, to generate a new consensus state. The state of the environment model is used to control the robot's motors, and to guide a selective sensing mode.

A system which is entirely dependent on a-priori information for its model construction is prune to problems of misinterpretations of sensory information. Such misinterpretations may result from the dynamic nature of an environment, temporary sensor failure or bad performance, and other unpredictable but expected situations. An autonomous system should thus have the means to protect itself by being able to question the accuracy of its model and build a new one if necessary. To overcome the problems of a totally 'model-trusting' system we have experimented with a neural network which is trained by back-propagation to identify desired features in the environment. The network uses a special 1-step feedback of its previous output state which provides the system with first order dynamic continuity, and improves its performance. A simulation of this system is run off-line, and preliminary results in which a model of a corridor is built strictly based on sensory information are presented.

ENVIRONMENT/WORLD MODELS

An intelligent autonomous system working in an unstructured, dynamic environment requires models for navigation, planning, object recognition, and internal process control. The model used by such an autonomous agent is usually referred to as the World Model (WM). The WM includes information about the work space, objects, properties of objects, relationships among them, events that can occur, and any other relevant information[9].

We distinguish between the WM which contains relatively fixed information and the Environment Model (EM) which contains more detailed, dynamic, and explicitly task-oriented information. The WM and EM have similarities to long-term and short-term memories in many systems[11], though there are some fundamental differences. The most important difference is the explicit attempt to capture 'locality' (temporal, spatial, and contextual) of information in the EM.

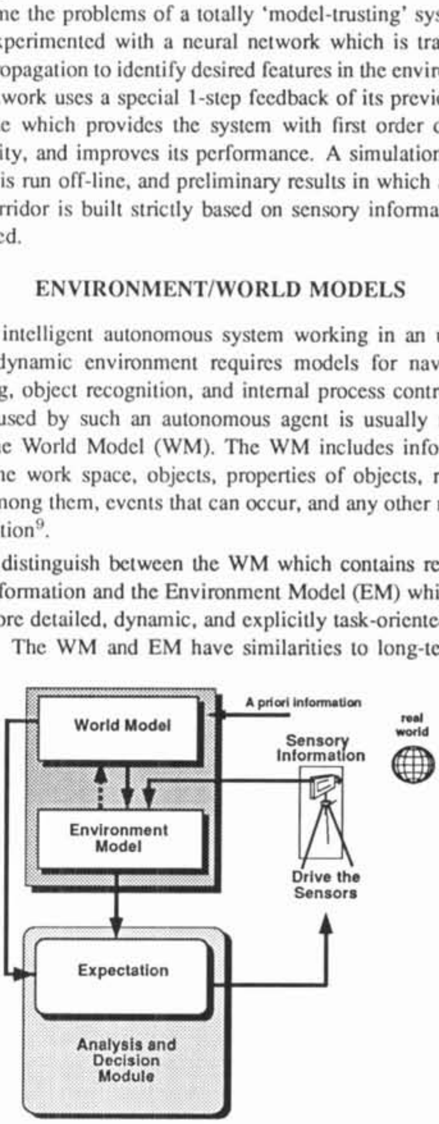

Figure 1: Information Flow Diagram

Our system demonstrates these concepts and some of the interactions between the models and the sensory information (see figure 1). A global data base – the World Model – represents all the information known a priori about the environment. In our test case (hall-following) this model is a floor plan of the hallway showing the exact position of each door and other openings along both walls. It includes the name of the room associated with each door. The environment model, on the other hand, contains dynamic information such as the robot's position and orientation, position of landmarks relative to the robot, state of objects in the environment etc. The environment model is created from a combination of sensory information and data from the World Model.

Finally, the information in the Environment Model and that from the World Model are combined to create the Expectation. The Expectation can be defined in terms of Environment Model structures (next position, orientation etc), or in terms of sensory data (actual expected readings from sensors). The Expectation is used to control the sensing direction and mode.

MODEL DRIVEN APPROACH

An autonomous agent interacting dynamically with the environment reacts to the sensed information and to changes in its internal model. The mechanisms that control these reactions can be divided into two main categories: sequential mechanisms, or event driven mechanisms. Sequential mechanisms follow a predetermined sequence of operations, whereas the event driven mechanisms will vary according to the changing events. The disadvantages of sequential mechanisms lie with their inadequacy to deal with general cases. Their advantage, however, is that bounds on the computation time can be easily derived. This is especially critical for meeting completion time requirements in real-time systems.

Event driven mechanisms can be further divided into two sub-categories: data driven, and model driven. In the data driven mechanisms sensory information is used to control the algorithms, in the model driven approach model information is used. The model driven mechanism uses processed information, thus takes into account the uncertainty in sensory information, previous data, and expectations. Also, in these mechanisms, the complexity of the operation is proportional to the complexity of the model, and not to that of the sensed data. Roth-Tabak and Jain[12] have demonstrated a model driven approach for information assimilation and model building using multiple views. This allows their algorithm to be fully parallelizable. Data driven mechanisms, on the other hand, can be used for reflexive behavior, when immediate reaction is necessary to avoid potentially dangerous situations, in which case information is extracted directly from the sensors to save time. The danger is that the system loses contact with the real environment because the extensive use of data driven mechanisms can generate 'confusions' due to bad data or wrong interpretation.

Our system combines all three mechanisms :

- A sequence of operations mainly sensing, integrating the information, and reacting, guarantees that although the robot is constantly in motion, it is also constantly 'in touch' with the environment.
- Simple input data analysis is performed repeatedly and guarantees that the robot is keeping away from 'troubles', namely avoiding obstacles, and keeping "on track" as it moves down the hall.
- Expectations are used to direct the sensors. This is a central mechanism since it mediates the flow of information back into the model, which is used to establish the updated environment model values. This is a focus of attention

mechanism which helps the system to ignore large pieces of potentially distracting information and concentrate its full analysis power on the critical information. It is the key to the real-time capability of our system.

Assimilating Uncertain Information

Information from the sensors is assimilated into the environment model on two different axes: 1. information from different sensors given at the same time, and 2. assimilation over time. Sensory information is never fully reliable, however, and one has to to be careful when computing average values, to filter out unreasonable values. In our system this is handled by defining flexible uncertainty ranges for the various variables in the environment model, and by generating anticipation values for the sensory readings based on time lapses and dead-reckoning between the sensing intervals.

The information in the environment model is represented as intervals of uncertainty. For example, the robot's longitude position in the hall is given as an interval of size proportional to the uncertainty of the information. As the robot moves and collects more sensory readings, these uncertainty ranges are updated using two basic schemes: *local updating*, and *global updating*.

In the local updating scheme different sensory information contributes to the updating process. As each data modality carries a different error margin, a new uncertainty range (or set of ranges) is computed for each by taking into consideration the largest error margins. The final range is the intersection of the 'acceptable' readings from all the modalities.

As the local updating is based on the previous instantiated model, although we are trying to use worst case measures, additive errors can creep into the model and confuse the system particularly in cases where the environment consists of repetitive structures. The global updating is introduced to avoid this problem. In this scheme a global uncertainty range is also kept. It is computed relative to an earlier position for which some uncertainty may exist, but for which the true value is known to be within the uncertainty range. The global uncertainty is updated using sensing modalities that can introduce only 'noisy' results but no outliers.

Expectation in our system is used not only to drive the sensors, but also to filter out improbable ('non acceptable') readings. The system uses the previous model and its knowledge of the dynamics, the elapsed time or dead-reckoning, and the World model to anticipate particular readings. The more readings the system accepts the more reliable the model becomes and the smaller the envelop of acceptable readings becomes. If the actual readings are outside of an acceptable envelop they are discarded, and the envelop margins grow.

EXPERIMENTAL RESULTS – USING A MODEL

Using the above guidelines a simple robot (Heathkit 2000) can move down a hall at a continuous pace while constantly collecting measurements correcting its orientation, determining its position in the hall, and recognizing special landmarks (doors and openings).

The world model which defines the hallway consists of a combination of metric and non-metric information. The model represents 'places' such as openings and doors as intervals. These places determine the sensing mode to be used for properly following the hall. In addition, landmarks such as door beginnings or endings are represented metrically. This information is used to correct the dead-reckoning and detect the actual position of the robot in the hallway (which is defined as an interval of size relative to the certainty of the information).

The environment model consists of the section of the hallway in the vicinity of the robot (properly oriented), and the robot's state. The robot's state consists of its current position (longitude and latitude uncertainty ranges and closest landmarks), current orientation, the latest filtered sensor readings, other parameters representing the system's interpretations of its latest sensory information, and sensing mode (based on the surrounding features).

Due to the inherent communication delays in the given system, and to our initial requirement of real-time operation, the sensing control has to be very selective. The system uses two sensing modes, in one it is trying to determine its latitude position and orientation in the hall (which is crucial for survival in the 'hall-jungle'), in the next the system is trying to recognize detectable features in the environment and match them to the model (see figure 2).

Our current system is able to navigate itself successfully in the given corridor using three levels of action:

- Reactive – the system fires the base sonar sensor constantly, and whenever the range is too small it stops and waits for the obstacle to be moved.
- Corrective – Orientation and distance from the walls are kept in a band of values which is large enough to let it move uninterrupted but is small enough to deal with the potential movements between the measurement cycles.
- Planned – Selective sensor readings are taken only from desired orientation, and at positions where, from the world and environment models, they are expected to yield useful results.

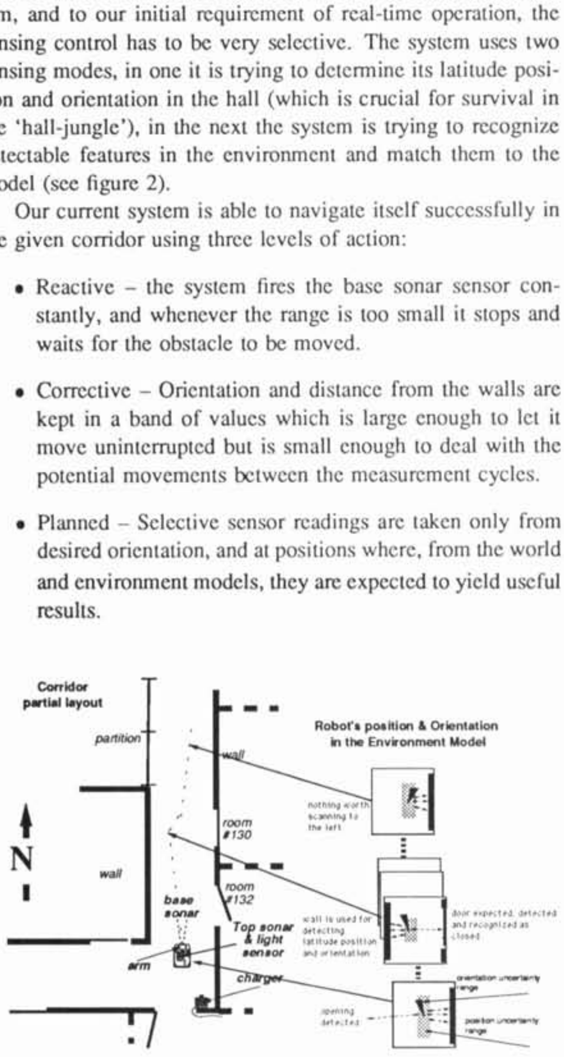

Figure 2: Layout for hall-following experiment

The robot ignores unexpected readings, such as created by people moving by it, so it is able to deal with this type of dynamics. The robot also recognizes hall features such as openings and doors, and is able to identify closed and open doors. We have also demonstrated the generality of the model and the system by picking a different corridor in the building, and simply loading the new world model of this other corridor with no additional changes to the system. The system was able to perform successfully.

GENERATING A MODEL FROM SENSORY DATA

The details of the actual techniques used to generate the model are beyond the scope of this paper. We shall just briefly outline the means and present preliminary results. The aim of this project was to investigate the use of a neural network to recognize features of interest in a corridor without any a priori knowledge (or model) of the environment. Input data consisted of local sonar and light measurements, and a trained neural net is used to identify the various features of interest in the vicinity of the robot. The system can then use this information to guide the robot along the corridor, and to build a model of the corridor. This information can be assimilated with previously available information, or be used to construct the model in a strictly data-based mode.

Figure 3 presents the results graphically. The actual corridor walls and doors are represented as the 3-D shaded structures. The model constructed by the neural net is shown as the dark thick lines parallel to the walls. Doors are represented as line indentations and open-doors as jagged indentations.

Except for a single gross error where the network hypothesized a door which was not there, the network performed extremely well: All the doors, openings and walls (even the single open door) were correctly identified. There was a varying amount of positional error for the identified features shown in the figure, but it was in the order of magnitude of the robot's single 'step size'. The figure should provide a reasonable idea of the amount of positional error, as it is scaled to represent the proportions correctly - except for the width of the corridor which was increased to make the figure intelligible. To get some idea of the actual dimensions, a door in reality is 35 inches wide.

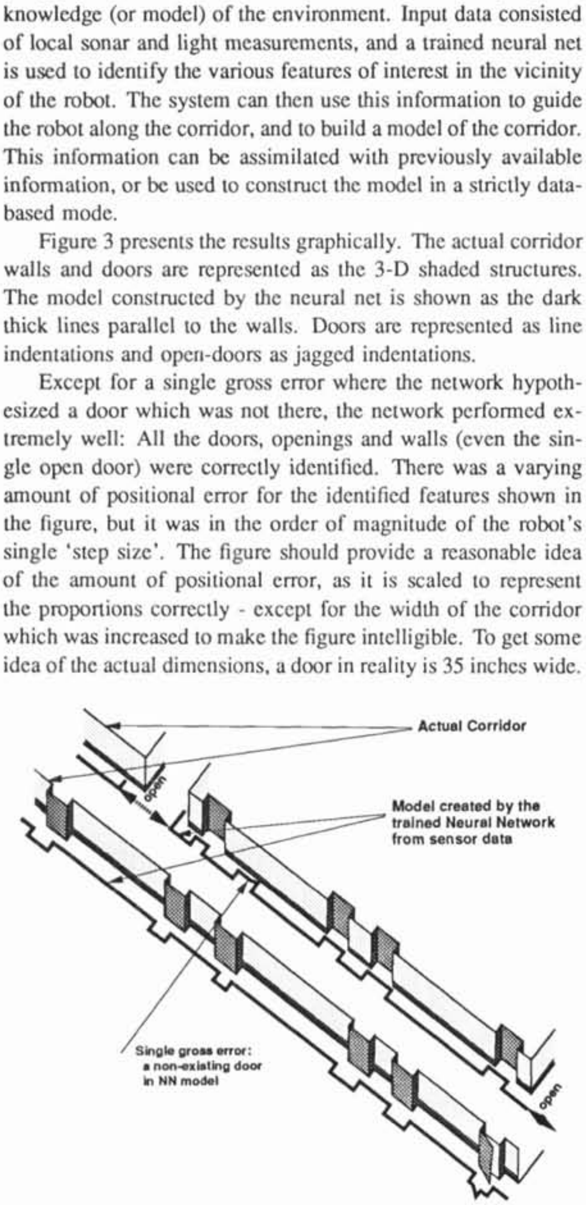

Figure 3: A corridor model generated by the Neural Net

CONCLUSION

In the presented method, we have created a simple world model representing a hallway using a combination of metric and qualitative information. We have demonstrated our capability of dealing with uncertainty in sensor data and robot position by driving the system using an environment model, and by generating anticipation for both updated values of environment model variables such as position and for actual sensor readings. Despite the slow communication between the off-board process and the controllers on the robot, the slow reaction of the onboard controllers through a BASIC interpreter, the slow sensing capability, the infamous reliability of the sonar sensors, and the sparse amount of available information, our system performs successfully in real-time.

The scheme demonstrated in this simple system can be extended to more complex systems and situations. More complex sensing devices will increase the potential 'vocabulary' the world model can handle since more features are detectable and possibly with greater accuracy.

We have also presented preliminary results of generating a model directly from sensory information. A combination of such a capability will render a system much more robust as it will be not only able to use a priori information, but also question and verify it, and produce a similar model for unknown environments.

REFERENCES

1. Nils Nilsson, "Shakey the Robot", Artificial Intelligence Center Technical Note 323, SRI International, Menlo Park, 1984.

2. Hans P. Moravec, Robot Rover Visual Navigation, UMI Research Press, Ann Arbor, Michigan, 1981.

3. Ernst Dieter Dickmanns and Volker Graefe, "Dynamic Monocular Machine Vision", *Machine Vision and Applications*, 1(4):223-240, 1988.

4. David J. Kriegman, Ernst Triendl, and Thomas O. Binford, "Stereo Vision and Navigation in Buildings for Mobile Robots", *IEEE Trans. Robotics and Automation*, 5(6):792–803, December 1989.

5. J. Borenstein and Y. Koren, "The Vector Field Histogram - Fast Obstacle Avoidance for Mobile Robots", *To be published as a paper in IEEE Journal of Robotics and Automation*, Revised 1990.

6. Rodney A. Brooks, "Intelligence Without Representation", In *Proceeding of the Workshop on the Foundations of Artificial Intelligence*, 1987.

7. Long-Ji Lin, Reid Simmons, and Christopher Fedor, "Experience with a Task Control Architecture for Mobile Robots", Tech Report CMU-RI-TR-89-29, The Robotics Institute, Carnegie Mellon U., December 1989.

8. Charles Thorpe and Takeo Kanade, "Carnegie Mellon Navlab Vision", In *Proc. DARPA Image Understanding Workshop*, pages 273–282, May 1989.

9. R. Jain, "Environment Models and Information Assimilation", Technical Report RJ6866, IBM Almaden Research Center, June 1989.

10. Benjamin J. Kuipers and Tod S. Levitt, "Navigation and Mapping in Large-Scale Space", *AI Magazine*, 9(2):25–43, Summer 1988.

11. Daryl T. Lawton, Tod S. Levitt, Christopher C. McConnell, Philip C. Nelson, and Jay Glicksman, "Environmental Modeling and Recognition for an Autonomous Land Vehicle", In *Proc. DARPA Image Understanding Workshop*, pages 107–121, 1987.

12. Yuval Roth-Tabak and Ramesh Jain, "Building an Environment Model Using Depth Information", *Computer*, 22(6):85–90, June 1989.