Vision-Based Map-Making and Navigation

Kurt Steinkraus & Leslie Kaelbling

Artificial Intelligence Laboratory
Massachusetts Institue Of Technology
Cambridge, Massachusetts 02139

http://www.ai.mit.edu

The Problem: We want to let a robot explore its environment, create a map of that environment, and then find its way around in it. We would like it to be able to do this using visual input. It should need as little human intervention, in the way of pre-supplied maps or other input, as possible. The system should also be robust with respect to sensor noise and sloppy actuators.

Motivation: Robots have a hard time finding their way around on their own. That's sometimes OK if there's a person close by controlling the robot, but in some situations, this isn't practical. A mobile robot being sent to another planet could not be guided remotely because of the large time-delay involved. A seeing-eye robot would exist specifically to guide a person through the world, rather than the other way around. A small army of robots moving around a factory floor, making and delivering things, would require an unreasonably large number of operators. An actual army of robots (hi DARPA!) might need to be able to navigate unmapped hostile territory and make decisions without there being a vulnerable communications link back to controllers.

The reason we will focus on vision, as opposed to sonar sensors or the like, is that vision has the potential to far outperform other methods, supplying much more information.

Previous Work: Some previous approaches (e.g. [1]) try to construct an global 3D geometric model of the world, and then to figure out how to drive a robot around using that model. This approach is not how people make maps and seems doomed to eventual failure. One reason is that the geometric model is difficult to obtain and make consistent, due to measurement errors, and it will change as objects shift slightly or as measurements are noisy. Another is that tiny errors in robot navigation can add up to confuse the robot about its current location and the location of objects close to it.

A lot of previous approaches have used sonar or laser sensors rather than vision. We believe, though, that a camera gives much more information than a sonar sensor or even a laser range-finder. A camera also allows locations to be more easily identified or distinguished. Using vision shifts the problem from that of having too little information to having too much.

Approach: We will approach this problem along the lines of [2]. We will build a topological map as a graph; the nodes will be junctions (i.e. places where there are choices about which way to go next) or interesting locations, and the edges will be paths between those locations. We plan to include approximate geometric information associated with nodes and edges, but this information will be local and will be used primarily to assist the robot when it visits that location.

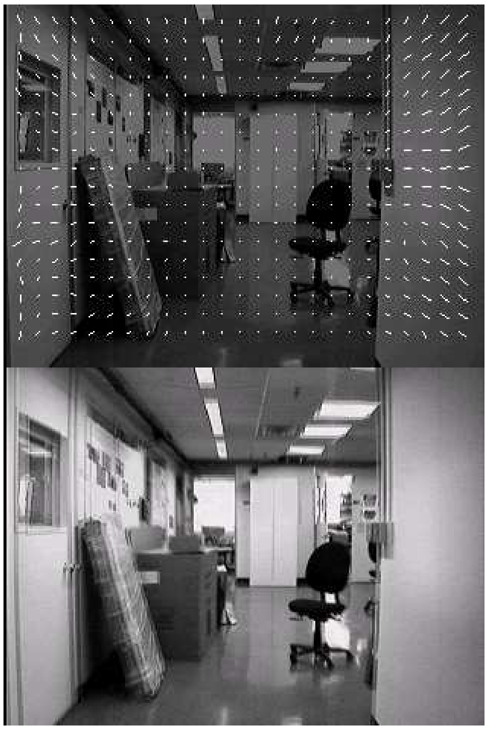
The work of another MIT grad student, Selim Temizer, will supply the robot's routines to maneuver around and recognize possible map features, such as doorways or T-junctions.

Difficulty: The hurdles are many, and I will just mention a few. Ones is that the robot needs to be able to recognize where it has been; this is sometimes difficult even for people, and is a nightmare for a robot because of image noise, moved objects, different lighting conditions, etc. Another is that it needs to be able to drive around without bumping into things and recognize when it has options about which way to go. Small obstacles may make a big room look like several small rooms, and differently sized passageways or doorways may be mistakenly identified or missed.

On top of all the difficulties in figuring out what the best way to do things is, the system should be able to operate (i.e. obtain mapping information and navigate about) in real time, which limits the possible complexity of algorithms used.

Impact: This work should allow robots to make useful maps of their environment, useful in that they can be used for navigation in an efficient, robust manner. Existing techniques for navigation in topological spaces can be applied to the maps learned. Having a robot map and navigate on its own has many industrial applications.

Future Work: Visual navigation has the potential to integrate with other abilities, such as fine motor control of a robotic arm or speech recognition and synthesis, to create a robot that acts much like a human. It would also be interesting to use other "senses," such as sonar, touch, or hearing, to assist map making.



Research Support: This work is supported by DARPA/ITO under contract number DABT 63-99-1-0012.

References:

[1] R. Chatila and J-P. Laumond. Position referencing and consistent world modeling for mobile robots. *IEEE International Conference on Robotics and Automation*, pp. 138-170, 1985.

[2] B. Kuipers. The Spatial Semantic Hierarchy.