

The Find-the-Remote Event

Ian Horswill

■ The Find-the-Remote event was considered the most challenging of the events in the 1997 AAAI Mobile Robot Competition and Exhibition. It required a broad range of both hardware and software capabilities. I discuss the rules and rationale for the event as well as the results.

The Find-the-Remote event was considered the hardest event of the contest. It involved fetching a known set of objects from unknown, but constrained, locations in a known environment. In real life, such functions might be useful for in-home care of the elderly or the physically disabled.

Why It Was Hard

This event was extremely difficult because it forced teams to implement both manipulation (the grasping and moving of objects) and visual object recognition. Furthermore, it explicitly required teams to implement them for a wide range of objects. It therefore eliminated a broad range of special-purpose sensing and manipulation strategies that would be specific to one or another class of objects. It also required that objects be lifted from a variety of surfaces (real furniture) at a variety of heights.

These requirements were enough to keep most teams from entering for simple hardware reasons. Many robots have insufficient on-board computing resources to perform real-time vision. Even fewer have real manipulators that can grasp objects at a distance. In addition, many robots aren't tall enough to even see an object on top of a kitchen table, much less grasp it. Thus, relatively few robots met the joint requirements of computing power, height, and manipulator work space.

Why It Was Important

Of course, it's the very difficulty of the task that makes it attractive. The consensus of the autonomous robot research community is that

we need to pursue realistic tasks in real, unmodified environments, particularly environments in which they must interact with humans. There is also broad agreement that the community needs to move on from tasks that exclusively involve locomotion to tasks that involve changing the environment, presumably through direct physical manipulation. Although previous AAAI robot competitions have involved some amount of manipulation, this competition event was the first that required dexterous manipulation of a range of objects along with the perceptual capabilities required to support it.

The Task

The rules specified a fixed course and a fixed set of objects that would populate it. The course consisted of typical household furniture and Lexan partitions arranged to produce a simplified two-room house. The objects were typical household objects, such as a television remote, a pill bottle, and fruits and vegetables. For obvious sanitary reasons, plastic models were used for the fruits and vegetables. Although the teams knew the possible objects and the layout of the course in advance, they did not know the locations of the objects. Therefore, robots had to perform a systematic search of all the flat surfaces of the environment to find the correct object. However, the locations of most objects were not completely arbitrary. Most objects were constrained to appear only in certain types of place. For example, fruits and vegetables would only be found in the kitchen, although they could be anywhere on any surface in the kitchen. Thus, the teams had the opportunity to use domain knowledge to optimize their robots' search routines.

The objects had a variety of appearances, requiring teams to use multiple visual cues, such as size, color, and texture, to classify them. Furthermore, the objects could appear on surfaces of different colors. To make the problem more manageable, however, the sur-

Object	Allowable Locations
Television remote	Television
Pill box	Coffee table or kitchen table
Coke can	Kitchen table or coffee table
Coffee cup	Anywhere
Cereal bowl	Kitchen table or sink
Videotape	Television or coffee table
Fruits and vegetables	Anywhere in kitchen
Ketchup bottle	Anywhere
Rubber chicken	Cutting board

Table 1. The Set of Objects.

faces were guaranteed to be textureless, and the objects were guaranteed to be well separated from one another. This approach simplified the visual problem of *segmentation* (determining which image region corresponds to which object) and prevented shape-recognition algorithms from having to worry about occlusion. It also simplified the problem of choosing grasp points by removing the issue of inadvertently grasping two objects.

The original intent had been for the tables, chairs, and other items of furniture to have known, and relatively distinctive, colors so that the robots could recognize them from a distance. However, this turned out to be difficult to arrange with the contractors. As it happened, though, no teams intended to use this feature. There were also a number of detailed rules for scoring and assigning partial credit in different contingencies. Because these rules were never invoked, I do not discuss them here.

The Course

The course consisted of a living room and a kitchen, partially separated by a Lexan partition. The rooms were populated with real living room furniture and the best available approximation to kitchen furniture. The furniture was obtained from a local contractor and was treated with near-religious reverence by

participants who feared that an accidental coffee stain might suddenly make them the unexpected owners of a \$2000 sofa bed in Providence, Rhode Island. A diagram of the course is shown in figure 1.

The Objects

The event used typical household objects. The particular set of objects was carefully chosen to require a collection of cues (color, texture, and shape) to distinguish them from one another. The set of objects is given in table 1.

Achieving the Task

To perform the task, a robot needed a number of capabilities: color-based object recognition and localization, texture-based object recognition, shape-based object recognition, grasp and pickup of a visually defined target, low-level locomotion control (for example, obstacle avoidance), spatial mapping and path planning, systematic search of a series of locations, and visual search of a given surface for an object.

Results

Initially, Kansas State University's (KSU) team (see sidebar) was the only group to enter the

Profile of a Winner: Kansas State University

The Kansas State University (KSU) robotics programming team won the Find-the-Remote event. The team's software was able to find, recognize, and retrieve all six items used in the preliminary round. Because there was no other competitor for the final round, it was turned into a demonstration with four items found and retrieved.

The team developed its winning software program on WILLIE, a NOMAD200 robot from Nomadic Technologies, Inc. WILLIE is a black cylinder approximately 2 feet in diameter and 3-feet tall, which weighs about 200 pounds. It is equipped with three wheels. On board the NOMAD is a PENTIUM computer with a hard drive. The NOMAD used in the competition is equipped with 2 rings of 16 sonar sensors. For the contest, an arm and a color camera were purchased. The LINUX operating system is used on board the robot. The competition software was written in C++.

The team's success was based on its software-engineering approach. In the requirements phase, the team identified some critical issues. These issues included learning to use the arm and the color camera, which were both new to the students. Algorithms for line, edge, and ellipse detection and camera calibration were investigated. The issues of mapping the environment, path planning, and robotic motion in the environment were familiar from class exercises as well as previous competitions.

The team needed a robust architecture for the robot. A layered architecture based on abstractions of the tasks was chosen. There were three levels in the object model: First, the bottom layer interfaced with the arm, motors, and sensors. Second, the middle layer contained item recognition, path planning, and item manipulation. Third, the top layer controlled the overall strategy. Each command reported success or failure to the calling method in the higher level. The calling method would retry the method or call an alternative method as appropriate. This approach allowed recovery from many errors, such as misalignment with the table or failure to pick up an already identified item.

The robot maintained a metric map of the environment. A number of locations were designated as viewing stations. Because each item was constrained to be in a few locations, the possible viewing stations for each item were associated



Courtesy, Karen Gustafson

WILLIE Positioned to Pick Up the Small Green Cup.

with the item. When the robot was trying to retrieve a specified item, it looked up the item and determined the nearest viewing station associated with the item. It then moved to the viewing station. Before looking for the item, it checked its position relative to the nearest table and compared it with the map to readjust its global position. If it found the item at the site, it aligned with the item, picked up the item, and returned to the starting location. If the item was not found, it moved to the next-closest viewing station associated with the item. If it exhausted all the viewing stations associated with an item without finding the item, it would go on to the next item on the list.

For use of the camera in recognition of the items, many approaches were investigated. An identification approach was determined for each item. The approach used color or simple size and shape, such as diameter of the ellipse at the top of the cup to distinguish the small cup from the large cup. Items were recognized in two distinct ways: First, many items were recognized by color. The recognition of these

items was done with HSV (hue, saturation, luminosity) thresholding and numerous filters to remove noise. Second, these thresholded images were then evaluated for a silhouette of the proper color and dimensions of the desired item. A cylinder, such as a cup, was recognized by finding the ellipse on the top of the cylinder. Once the ellipse was detected, the three-dimensional (3D) space mapping was used to find the exact diameter of the cylinder and, thus, a good indication of the identity of the item.

Because the camera and the arm are on opposite sides of the robot, the issue of positioning the robot precisely so that the robot can turn and pick up an item was important. An edge-detection algorithm was written using Lowe's algorithm (Rosin and West 1995) to detect the major edge of the table. The camera was calibrated (Pratt 1991), and trigonometry was used to create a mapping from any (x, y) point in a camera image whose height is known to its corresponding (x, y, z) position in 3D space. The exact position of the camera was critical for the algorithm, and it took about two hours to recalibrate the camera every time it was moved.

The height of each table was then used to determine the position of the table edge in 3D space relative to the robot. The 3D space mapping was also used to determine the exact point in 3D space of the item so that the arm could locate and retrieve the item without having to take additional images when approaching the item. Then the robot could be positioned to pick up the item (see the figure) with an accuracy of about a half inch.

At the contest site, the actual map for the contest environment was made; adjustments were made for lighting and, in particular, for glare on the tables from the overhead lights; and the vision-recognition routines were specialized for the actual items.

The KSU team for the 1997 AAAI Mobile Robot Competition and Exhibition consisted of Mike Novak, Todd Prater, Brian Rectanus, and Steve Gustafson. David Gustafson was the adviser.

– David Gustafson

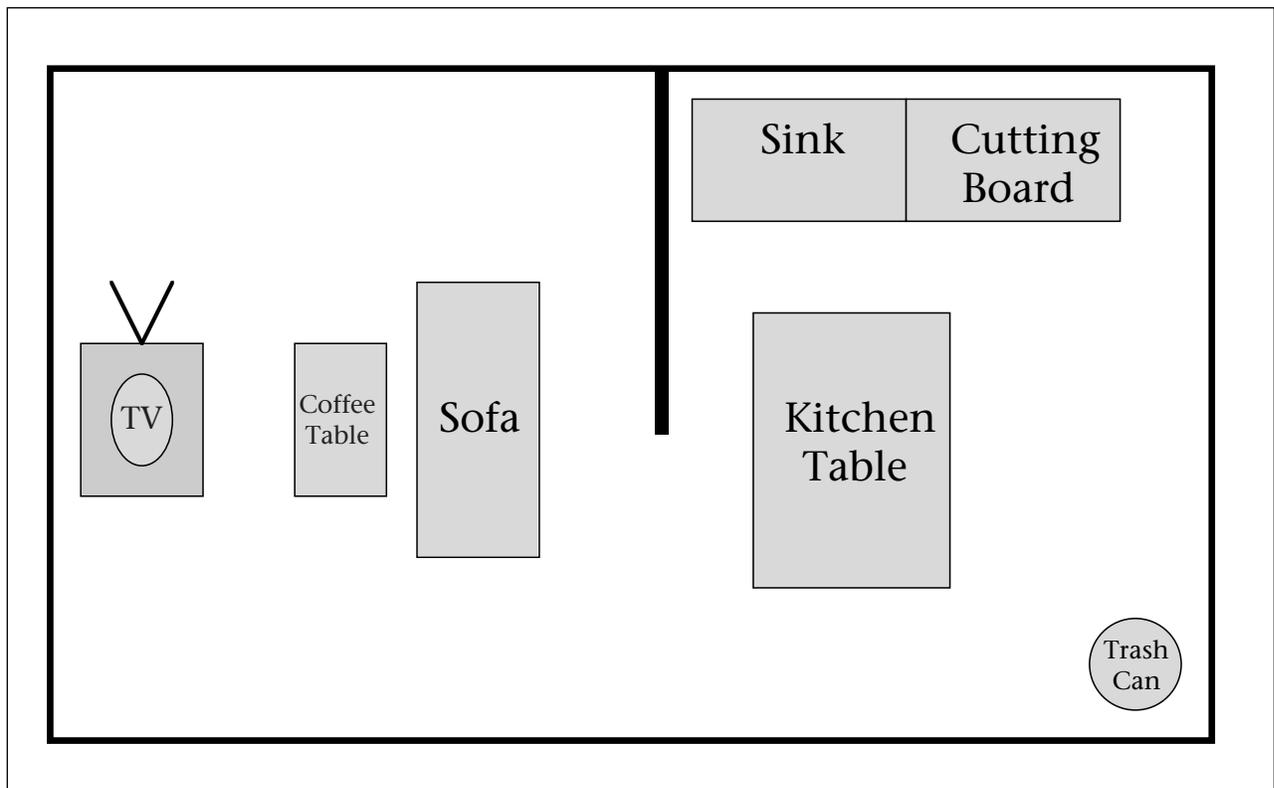


Figure 1. The Course.

event. The team put in many months of work and arrived at the competition with a working entry. Two other teams arrived at the competition with late entries but had to withdraw on the last day because of hardware problems.

Because of the lack of entries and the perceived difficulty of the event, a number of rules intended to complicate the task, or to ensure fair scoring, were abandoned. For example, under the original rules, teams didn't have direct access to the test objects that would be used in the competition (although online images would be available). This rule was meant, in part, to remove the possibility that one team might gain an unfair advantage by managing to find a hot-pink television remote that could be distinguished from other objects purely on the basis of color. However, with only one team, the rule became an unnecessary hassle, so the KSU team was allowed to bring its own objects for the competition.

In the end, the event was a success. The KSU team's robot successfully fetched and delivered all the objects that it was assigned.

Acknowledgments

The original draft rules were written by Jim Firby and Ron Arkin. The final version was writ-

ten by the event rules committee (Ian Horswill, Daniela Rus, and Robin Murphy).

References

- Pratt, W. K. 1991. *Digital Image Processing*. New York: Wiley.
- Rosin, P. L., and West, G. A. W. 1995. Nonparametric Segmentation of Curves into Various Representations. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 17:1140-1153.



Ian Horswill is an assistant professor of computer science at Northwestern University. He received his Ph.D. in 1993 from the Massachusetts Institute of Technology. His research focuses on the integration of sensory-motor and cognitive systems in autonomous robots.



David Gustafson is a professor in the Computing and Information Science Department at Kansas State University. His e-mail address is dag@cis.ksu.edu.