Auton Robot (2006) 21:29–41
DOI 10.1007/s10514-006-7223-8

Robot-assisted wayfinding for the visually impaired in structured indoor environments

**Vladimir Kulyukin · Chaitanya Gharpure ·
John Nicholson · Grayson Osborne**

Published online: 15 June 2006

Abstract We present a robot-assisted wayfinding system for the visually impaired in structured indoor environments. The system consists of a mobile robotic guide and small passive RFID sensors embedded in the environment. The system is intended for use in indoor environments, such as office buildings, supermarkets and airports. We describe how the system was deployed in two indoor environments and evaluated by visually impaired participants in a series of pilot experiments. We analyze the system's successes and failures and outline our plans for future research and development.

Keywords Assistive robotics · Robot-assisted wayfinding · RFID-based localization · Human-robot interaction

1. Introduction

Since the adoption of the Americans with Disabilities Act of 1990 that provided legal incentives for improvement in universal access, most of the research and development (R&D) has focused on removing *structural* barriers to universal access, e.g., retrofitting vehicles for wheelchair access, building ramps and bus lifts, improving wheelchair controls, and providing access to various devices through specialized interfaces, e.g., sip and puff, haptic, and Braille.

For the 11.4 million visually impaired people in the United States (LaPlante and Carlson, 2000), this R&D has done little to remove the main *functional* barrier: the inability to navigate dynamic and complex environments. This inability denies the visually impaired equal access to many private and public buildings, limits their use of public transportation, and makes the visually impaired a group with one of the highest unemployment rates (74%) (LaPlante and Carlson, 2000). Thus, there is a significant need for systems that improve the wayfinding abilities of the visually impaired, especially in unfamiliar environments, where conventional aids, such as white canes and guide dogs, are of limited use. In the remainder of this article, the term *unfamiliar* is used only with respect to visually impaired individuals.

1.1. Assisted navigation

Over the past three decades, considerable R&D effort has been dedicated to navigation devices for the visually impaired. Benjamin et al. (1973) built the C-5 Laser Cane. The cane uses optical triangulation with three laser diodes and three photo-diodes as receivers. Bissit and Heyes (1980) developed the Nottingham Obstacle Detector (NOD), a handheld sonar device that gives the user auditory feedback with eight discrete levels. Shoval et al. (1994) developed the NavBelt, an obstacle avoidance wearable device equipped with ultrasonic sensors and a wearable computer. The NavBelt produces a 120-degree wide view ahead of the user. The view is translated into stereophonic audio directions. Borenstein and Ulrich (1994) built GuideCane, a mobile obstacle avoidance device for the visually impaired. GuideCane consists of a long handle and a ring of ultrasonic sensors mounted on a steerable two-wheel axle.

V. Kulyukin (✉) · C. Gharpure · J. Nicholson
Computer Science Assistive Technology Laboratory (CSATL),
Department of Computer Science,
Utah State University, Logan, UT 84322-4205
e-mail: vladimir.kulyukin@usu.edu
e-mail: cpg@cc.usu.edu
e-mail: jnicholson@cc.usu.edu

G. Osborne
Department of Psychology, Utah State University, Logan, UT
84322 e-mail: grayson.osborne@usu.edu

More recently, a radio frequency identification (RFID) navigation system for indoor environments was developed at the Atlanta VA Rehabilitation Research and Development Center (Ross, 2001; Ross and Blasch, 2002). In this system, the blind users' canes are equipped with RFID receivers, while RFID transmitters are placed at hallway intersections. As the users pass through transmitters, they hear over their headsets commands like *turn left*, *turn right*, and *go straight*. The Haptica Corporation has developed Guido©, a robotic walking frame for people with impaired vision and reduced mobility (www.haptica.com). Guido©uses the onboard sonars to scan the immediate environment for obstacles and communicates detected obstacles to the user via speech synthesis.

While the existing approaches to assisted navigation have shown promise, they have had limited success for the following reasons. First, many existing systems increase the user's navigation-related physical load, because they require that the user wear additional and, oftentimes substantial, body gear (Shoval et al., 1994), which contributes to physical fatigue. The solutions that attempt to minimize body gear, e.g., the C-5 Laser Cane (Benjamin et al., 1973) and the GuideCane Borenstein and Ulrich (1994), require that the user effectively abandon her conventional navigation aid, e.g., a white cane or a guide dog, which is not acceptable to many visually impaired individuals. Second, the user's navigation-related cognitive load remains high, because the user makes all final wayfinding decisions. While device-assisted navigation enables visually impaired individuals to avoid immediate obstacles and gives them simple directional hints, it provides little improvement in wayfinding over white canes and guide dogs. Limited communication capabilities also contribute to the high cognitive load. Finally, few assisted navigation technologies are deployed and evaluated in their target environments over extended time periods. This lack of deployment and evaluation makes it difficult for assistive technology (AT) practitioners to compare different solutions and choose the one that best fits the needs of a specific individual.

1.2. Robotic guides

The idea of robotic guides is by no means novel. Horswill (1993) used the situated activity theory to build Polly, a mobile robot guide for the MIT AI Lab. Polly used lightweight vision routines that depended on textures specific to the lab.

Thrun et al. (1999) built MINERVA, an autonomous tour guide robot that was deployed in the National Museum of American History in Washington, D.C. The objective of the MINERVA project was to build a robot capable of educating and entertaining people in public places. MINERVA is based on Markov localization and uses ceiling mosaics as its main environmental cues. Burgard et al. (1999) developed RHINO, a close sibling of MINERVA, which was deployed as an interactive tour guide in the Deutsches Museum in Bonn, Germany. The probabilistic techniques for acting under uncertainty that were used in RHINO and MINERVA were later used in Pearl, a robotic guide for the elderly with cognitive and motor disabilities, developed by Montemerlo et al. (2002).

Unfortunately, these robotic guides do not address the needs of the visually impaired. The robots depend on the users' ability to maintain visual contact with them, which cannot be assumed for the visually impaired. Polly has very limited interaction capabilities: the only way users can interact with the system is by tapping their feet. To request a museum tour from RHINO (Burgard et al., 1999), the user must identify and press a button of a specific color on the robot's panel. Pearl also assumes that the elderly people interacting with it do not have visual impairments.

The approach on which Polly is based requires that a robot be evolved by its designer to fit its environment not only in terms of software, but also in terms of hardware. This makes it difficult to produce replicable solutions that work out of the box in a variety of environments. Autonomous solutions like RHINO, MINERVA, and Pearl also require substantial investments in customized engineering and training to become and, more importantly, to remain operational. While the software and hardware concerns may be alleviated as more on-board computer power becomes available with time, collisions remain a concern (Burgard et al., 1999).

Mori and Kotani (1998) developed HARUNOBU-6, a robotic travel aid to guide the visually impaired on the Yamanashi University campus in Japan. HARUNOBU-6 is a motorized wheelchair equipped with a vision system, sonars, a differential GPS, and a portable GIS. Whereas the wheelchair is superior to the guide dog in its knowledge of the environment, as the experiments run by the HARUNOBU-6 research team demonstrate, the wheelchair is inferior to the guide dog in mobility and obstacle avoidance. The major source of problems was vision-based navigation, because the recognition of patterns and landmarks was greatly influenced by the time of day, weather, and season. In addition, the wheelchair is a highly customized piece of equipment, which negatively affects its portability across a broad spectrum of environments.

1.3. Robot-assisted wayfinding

Any R&D endeavor starts with the basic question: is it worthy of time and effort? We believe that with respect to robot-assisted wayfinding for the visually impaired this question can be answered in the affirmative. We offer several reasons to justify our belief. First, robot-assisted wayfinding offers feasible solutions to two hard problems perennial to wearable assisted navigation devices for the visually impaired: hardware miniaturization and portable power supply. The amount of body gear carried by the user is significantly minimized, because most of it can be mounted on the robot and powered from on-board batteries. Therefore, the navigation-related physical load is reduced. Second, since such key wayfinding capabilities as localization and navigation are delegated to the robotic guide, the user is no longer responsible for making all navigation decisions and, as a consequence, can enjoy a smaller cognitive load. Third, the robot can interact with other people in the environment, e.g., ask them to yield or receive instructions. Fourth, robotic guides can carry useful payloads, e.g., suitcases and grocery bags. Finally, the user can use robotic guides in conjunction with her conventional navigation aids, e.g., white canes and guide dogs.

In the remainder of this article, we will argue that robot-assisted wayfinding is a viable universal access strategy in structured indoor environments where the visually impaired face wayfinding barriers. We begin, in Section 2, with an ontology of environments that helps one analyze their suitability for robot-assisted wayfinding. In Section 3, we describe our robotic guide for the visually impaired. We specify the scope limitations of our project and present the hardware and software solutions implemented in the robotic guide. Section 4 discusses robot-assisted wayfinding and the instrumentation of environments. In Section 5, we describe the pilot experiments conducted with and without visually impaired participants in two structured indoor environments. We analyze our successes and failures and outline several directions for future R&D. Section 6 contains our conclusions.

2. An ontology of environments

Our ability to operate in a given environment depends on our familiarity with that environment and the environment's complexity (Tinbergen, 1976; Agre, 1988). When we began our work on the robotic guide, we soon found ourselves at a loss as to what criteria to use in selecting target environments. This lack of analytical framework caused us to seek an *operational* ontology of environments. After conducting informal interviews with visually impaired individuals on environmental accessibility and analyzing system deployment options available to us at the time, we decided to classify environments in terms of their *familiarity* to the user, their *complexity*, and their *containment*.

In terms of user familiarity, the ontology distinguishes three types of environments: *continuous*, *recurrent*, and *transient*. Continuous environments are environments with which the user is closely familiar, because she continuously interacts with them. For example, the office space in the building where the user works is a continuous environment. Recurrent environments are environments with which the user has contact on a recurrent but infrequent basis, e.g., a conference center where the user goes once a year or an airport where the user lands once or twice a year. Recurrent environments may undergo significant changes from visit to visit and the user may forget most of the environment's topology between visits. Transient environments are environments with which the user has had no previous acquaintance, e.g., a supermarket or an airport the user visits for the first time.

Two types of environmental complexity are distinguished in the ontology: *structural* and *agent-based*. Structural complexity refers to the physical layout and organization of a given environment, e.g., the number of halls, offices, and elevators, the number of turns on a route from A to B, and the length of a route from A to B. Agent-based complexity refers to the complexity caused by other agents acting in the environment, and is defined in terms of the number of static and dynamic obstacles, e.g., boxes, pieces of furniture, and closed doors, and the number of people en route.

Our ontology describes environmental complexity in terms of two discrete values: *simple* and *complex*. Hence, in terms of environmental complexity, the ontology distinguishes four types of environments: (1) simple structural, simple agent-based; (2) simple structural, complex agent-based; (3) complex structural, simple agent-based; and (4) complex structural, complex agent-based. It should be noted that, in terms of its agent-based complexity, the same environment can be simple and complex at different times. For example, the agent-based complexity of a supermarket at 6:00 am on Monday is likely to be much less complex than at 11:00 am on Saturday. Similarly, the agent-based complexity of a student center at a university campus changes significantly, depending on whether or not the school is in session.

In terms of containment, the ontology distinguishes two types of environment: *indoor* and *outdoor*. Thus, our ontology distinguishes a total of eight environments: the four above types classified according to environmental complexity, each of which can be either indoor or outdoor.

Given this ontology, we proceed to the next basic question: are all environments suitable for robot-assisted wayfinding? We do not think so. There is little need for such guides in *continuous* environments, i.e., environments with which the user is very familiar. As experience shows (Pfaffenberger et al., 1976), conventional navigation aids, such as white canes and guide dogs, are quite adequate in these environments, because either the user or the user's guide dog has an accurate topological map of the environment.

We do not think that robotic guides are suitable for outdoor environments either. The reason is twofold. First, outdoor environments are not currently within reach of robots unless the robots are teleoperated, at least part of the time (Fong and Thorpe, 2001). To put it differently, the state of the art in outdoor robot navigation technology does not yet allow one to reliably navigate outdoor environments. Second, the expense of deploying and maintaining such systems may be prohibitive not only to individuals, but to many organizations as well. Naturally, as more progress is made in vision-based outdoor navigation, this outlook is likely to change.

We believe that recurrent or transient indoor environments, e.g., supermarkets, airports, and conference centers, are both feasible and socially valid for robot-assisted navigation (Kulyukin et al., 2005). Guide dogs, white canes, and other navigation devices are of limited use in such environments, because they cannot help their users localize and find paths to useful destinations. Furthermore, as we argue below, such environments can be instrumented with small sensors that make robot-assisted wayfinding feasible.

3. RG-I: A robotic guide

In May 2003, the Computer Science Assistive Technology Laboratory (CSATL) at the Department of Computer Science of Utah State University (USU) and the USU Center for Persons with Disabilities (CPD) launched a collaborative project whose objective is to build an indoor robotic guide for the visually impaired. We have so far built, deployed and tested one prototype in two indoor environments. Our guide's name is RG-I, where "RG" stands for "robotic guide." We refer to the approach behind RG-I as *non-intrusive instrumentation of environments*. Our basic objective is to alleviate localization and navigation problems of completely autonomous approaches by instrumenting environments with inexpensive and reliable sensors that can be placed in and out of environments without disrupting any indigenous activities (Kulyukin and Blair, 2003). Additional requirements are: (1) that the instrumentation be reasonably fast and require only commercial off-the-shelf (COTS) hardware components; (2) that sensors be inexpensive, reliable, easy to maintain (no external power supply), and provide accurate localization; (3) that all computation run onboard the robot; and (4) that human-robot interaction be both reliable and intuitive from the perspective of the visually impaired users.

3.1. Scope limitations

Several important issues are beyond the scope of our project. First, robotic guides prototyped by RG-I are not meant for individual ownership. Rather, we expect institutions, e.g., airports, supermarkets, conference centers, and hospitals, to operate such guides on their premises in the future. One should think of RG-I as a step toward developing robotic navigational redcap services for the visually impaired in airport-like environments.

Second, it is important to emphasize that robotic wayfinding assistants prototyped by RG-I are not intended as replacements for guide dogs. Rather, these service robots are designed to complement and enhance the macro-navigational performance of guide dogs in the environments that are not familiar to the guide dogs and/or their handlers.

Third, we do not address the issue of navigating large open spaces, e.g., large foyers in hotels. While some references in the localization literature suggest that ultrasonic sensors could be used to address this issue (Addlesee et al., 2001), the proposed solutions are sketchy, have been deployed in small, carefully controlled lab environments, and do not yet satisfy the COTS hardware requirement. In addition, the ultrasonic sensors used in these evaluations must have external power sources, which makes both maintenance and deployment significantly harder. Thus, we currently assume that all environments in which RG-I operates are structured indoor environments, i.e., have walls, hallways, aisles, rows of chairs, T- and X-intersections, and solid and static objects, e.g., vending machines and water fountains, that the robot's onboard sensors can detect.

3.2. Hardware

RG-I is built on top of the Pioneer 2DX commercial robotic platform (See Fig. 1) from the ActivMedia Corporation. The platform has three wheels, two drive wheels in the front and a steering wheel in the back, and is equipped with three rechargeable Power Sonic PS-1270 onboard batteries. What turns the platform into a robotic guide is a Wayfinding Toolkit (WT) mounted on top of the platform and powered from the on-board batteries. As shown in Fig. 1, the WT resides in a polyvinyl chloride (PVC) pipe structure attached to the top of the platform. The WT includes a Dell Ultralight X300 laptop connected to the plat-

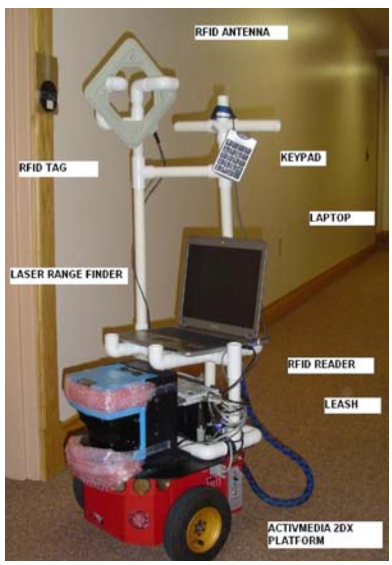


Fig. 1 RG-I: A robotic guide

form's microcontroller, a SICK LMS laser range finder from SICK, Inc., and a TI Series 2000 radio-frequency identification (RFID) reader from Texas Instruments, Inc.

The laptop interfaces to the RFID reader through a usb-to-serial cable. The reader is connected to a square 200 mm by 200 mm RFID RI-ANT-GO2E antenna that detects RFID sensors (tags) placed in the environment. Figure 2 shows several TI RFID Slim Disk tags. These are the only types of tags currently used by the system. These tags can be attached to any objects in the environment or worn on clothing. They do not require any external power source or direct line of sight to be detected by the RFID reader. They are activated by the spherical electromagnetic field generated by the RFID antenna with a radius of approximately 1.5 meters.

Several research efforts in mobile robotics have also used RFID technology in robot navigation. Kantor and Singh used RFID tags for robot localization and mapping (Kantor and Singh, 2002). Once the positions of the RFID tags are known, their system uses time-of-arrival information to estimate the distance from detected tags. Tsukiyama (2003) developed a navigation system for mobile robots using RFID tags. The system assumes perfect signal reception and measurement. Hähnel et al. (2003) developed a robotic mapping and localization system to analyze whether RFID can be used to improve the localization of mobile robots in office environments. They proposed a probabilistic measurement model for RFID readers that accurately localizes RFID tags in a simple office environment.

3.3. Software

RG-I has a modular hybrid architecture that consists of three main components: a path planner, a behavior manager, and a user interface (UI). The UI has two input modes: haptic and speech. The haptic mode uses inputs from a hand-held keypad; the speech mode accepts inputs from a wireless wearable microphone coupled to Microsoft's SAPI 5.1 speech recognition engine. The UI's output mode uses non-verbal audio beacons and speech synthesis.

a. RFID tag at a door b. RFID tag at a turn

Fig. 2 Deployed RFID tags

The UI and the planner interact with the behavior manager through socket communication. The planner provides the robot with path plans from start tags to destination tags on demand. The behavior manager executes the plans and detects plan execution failures.

This architecture is inspired by and partially realizes Kupiers' Spatial Semantic Hierarchy (SSH) (Kupiers, 2000). The SSH is a framework for representing spatial knowledge. It divides spatial knowledge of autonomous agents, e.g., humans, animals, and robots, into four levels: the control level, causal level, topological level, and metric level. The control level consists of low level mobility laws, e.g., trajectory following and aligning with a surface. The causal level represents the world in terms of views and actions. A view is a collection of data items that an agent gathers from its sensors. Actions move agents from view to view. For example, a robot can go from one end of a hallway (start view) to the other end of the hallway (end view). The topological level represents the world's connectivity, i.e., how different locations are connected. The metric level adds distances between locations.

The path planner realizes the causal and topological levels of the SSH. It contains the declarative knowledge of the environment and uses that knowledge to generate paths from point to point. The behavior manager realizes the control and causal levels of the SSH. Thus, the causal level is distributed between the path planner and the behavior manager.

The control level is implemented with the following low-level behaviors all of which run on the WT laptop: *follow-hallway*, *turn-left*, *turn-right* , *avoid-obstacles*, *go-thru-doorway*, *pass-doorway*, and *make-u-turn*. These behaviors are written in the behavior programming language of the ActivMedia Robotics Interface for Applications (ARIA) system from ActivMedia Robotics, Inc. Further details on how these behaviors are implemented can be found in Kulyukin et al. (2004) and Gharpure (2004).

The behavior manager keeps track of the global state of the robot. The global state is shared by all the modules. It holds the latest sensor values, which include the laser range finder readings, the latest detected RFID tag, current velocity, current behavior state, and battery voltage. Other state parameters include: the destination, the command queue, the plan to reach the destination, and internal timers.

4. Wayfinding

Visually impaired individuals follow RG-I by holding onto a dog leash. The leash is attached to the battery bay handle on the back of the platform. The upper end of the leash is hung on a PVC pole next to the RFID antenna's pole. Figure 3

Fig. 3 RG-I leading a guide dog user

shows a visually impaired guide dog user following RG-I. RG-I always moves closer to the right wall to follow the flow of traffic in structured indoor environments.

4.1. Instrumentation of environments

In instrumenting indoor environments with RFID tags, the following four guidelines are followed:

1. Every tag in the environment is programmed with a unique ID and placed on a non-metallic padding to isolate it from metallic substances in the walls;
2. Every door is designated with a tag;
3. Every object in the environment that can serve as a destination, e.g., a soda machine or a water fountain, is also tagged;
4. Every turn is designated with a tag; the tags are placed about a meter away from each turn.

After the tags are deployed in the environment, the knowledge base of the environment is engineered manually. The knowledge base represents an aerial view of the environment in which RG-I operates. The knowledge base for an environment consists of a tag connectivity graph, tag to destination mappings, and low-level behavior scripts associated with specific tags. Figure 4 shows a subgraph of the connectivity graph used in RG-I. The path is a *tag-behavior-tag* sequence. In the graph, *f*, *u*, *l* and *r* denote *follow-hallway*, *make-u-turn*, *turn-left*, and *turn-right*, respectively.

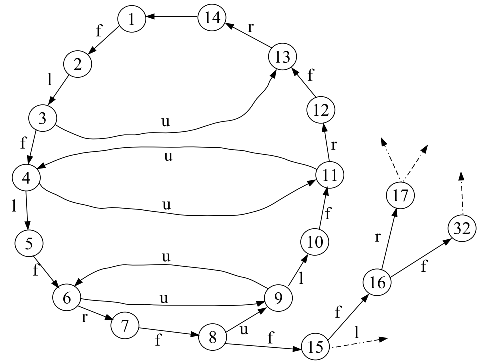


Fig. 4 A subgraph of a connectivity graph used in RG-I

The planner uses the standard breadth first search (BFS) to find a path from the start tag to the destination tag. For example, if the start tag is 4 and the destination tag is 17, *(4 l 5 f 6 r 7 f 8 f 15 f 16 r 17)* is a path plan. The plan is executed as follows. The robot detects tag 4, executes a left turn until it detects tag 5, follows the hallway until it detects tag 6, executes a right turn until it detects tag 7, follows the hallway until it detects tag 8, follows the hallway until it detects tag 15, follows the hallway until it detects tag 16, and executes a right turn until it detects tag 17. Given the tag connectivity graph, there are only two ways the robot can localize the user in the environment: (1) the user is approaching X, where X is the location tagged by the next tag on the robot's current path; and (2) the user is at X, where X is the location tagged by the tag that is currently visible by the robot's antenna.

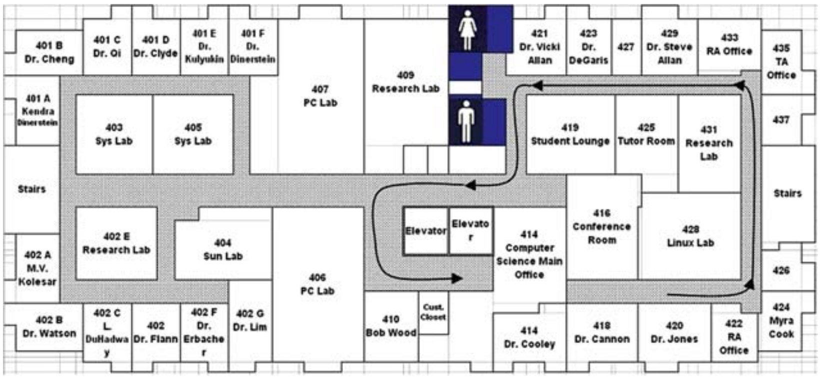
As another example, consider Fig. 5. The figure shows a map of the USU CS Department with a route that the robot is to follow. Figure 6 shows a path returned by the planner. Figure 7 shows how this path projected on the map of the environment as it is followed by the robot. Figure 7 also shows how the robot switches from one navigational behavior to another as its RFID antenna detects the tags on the path.

4.2. Obstacle avoidance

Obstacle avoidance is critical to robots navigating dynamic, structurally complex environments. Over the past two decades several obstacle avoidance techniques have been developed and tried on mobile robots in a variety of environments. The most prominent of those techniques are the potential fields approach (PF) (Khatib, 1985), the dynamic window approach (μdWA) (Burgard et al., 1999), the vector field histogram (VFH) (Borenstein and Koren, 1989), and the curvature velocity method (CVM) (Simmons, 1996).

While navigation in RG-I utilizes PF techniques, the overall approach to navigation implemented in RG-I differs from the above approaches in three respects. First, our approach

Fig. 5 USU CS Department



does not focus on *interpreting* obstacles detected by sensors and generating motion commands as a result of that interpretation. Instead, we focus on *empty spaces* that define the navigational landmarks of many indoor environments: hallways, turns, X- and T-intersections, etc. The robot itself cannot interpret these landmarks but can navigate them by following paths induced by empty spaces. Second, navigation in RG-I is not egocentric: it is distributed between the robot and the environment insomuch as the environment is instrumented with sensors that assist the robot in its navigation. Third, navigation in RG-I is *orientation-free*, i.e., the robot's sensor suite does not include any orientation sensor, such as a digital compass or an inertia cube; nor does the robot infer its orientation from external signals through triangulation or trilateration.

The robot's PF is a 10×30 egocentric grid. Each cell in the grid is $200\,\text{mm} \times 200\,\text{mm}$. The grid covers an area of 12 square meters (2 meters in front and 3 meters on each side). The grid is updated continuously with each scan of the laser range finder. A 180° laser scan is taken in front of the robot. The scan consists of a total of 90 laser range finder readings, taken at every 2 degrees. A laser scan is taken every 50 milliseconds, which is the length of time of an average action cycle in the ARIA task manager.

The exact navigation algorithm for determining the direction of travel executed by RG-I is as follows:

1. Do the front laser scan.
2. Classify each cell in the grid as free, occupied, or unknown and assign directions and magnitudes to the vectors in occupied cells.
3. Determine the maximum empty space.
4. Assign directions to the vectors in free cells.

The robot's desired direction of travel is always in the middle of the *maximum empty space*, a sector of empty space in front of the robot. Further details on the local navigation algorithms used in RG-I can be found in (Kulyukin et al., 2004) and (Gharpure, 2004).

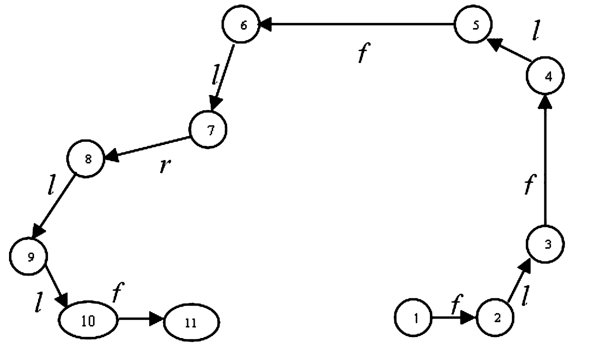


Fig. 6 A path of RFID tags and behaviors

4.3. Dealing with losses

We distinguish two types of losses: *recoverable* and *irrecoverable*. A recoverable loss occurs when the robot veers from a given path but reaches the destination nonetheless. In graph-theoretic terms, a veering event means that the original path is replaced with a different path. An irrecoverable loss occurs when the robot fails to reach the destination regardless of how much time the robot is given.

As shown in Fig. 8, there are two situations in which RG-I gets lost: 1) failure to determine the correct direction of travel and 2) RFID malfunction. The first situation occurs when the robot, due to its current orientation, finds the maximum empty space that causes it to veer from the correct path. In Fig. 8(a), RG-I detects the turn tag and, as prescribed by the plan, first executes the left turn behavior and then moves in the desired direction. However, since the hallway is blocked, RG-I veers away from the correct path. In Fig. 8(b), the turn tag is blocked by obstacles, which triggers the obstacle avoidance behavior. While avoiding the obstacle, the robot fails to detect the turn tag, because the tag falls outside the range of the RFID antenna, and does not make the left turn. The second situation that causes a loss, the RFID reader's malfunction, arises when the reader misreads the tag's ID or fails to activate the tag due to some interference in the environment. In our target environments, the second situation was rare.

Fig. 7 Run-time behavior switches

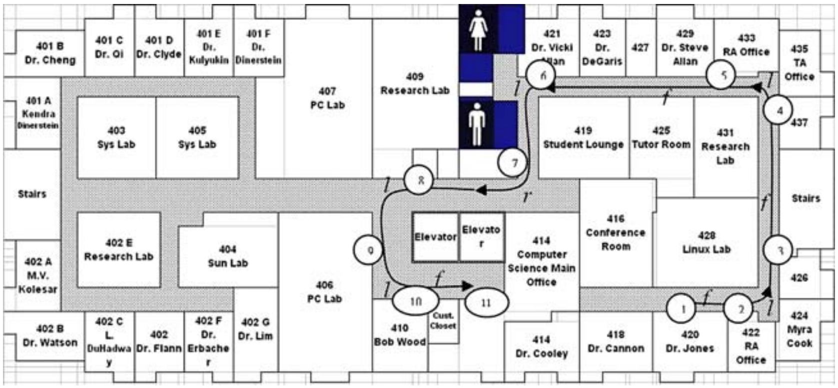
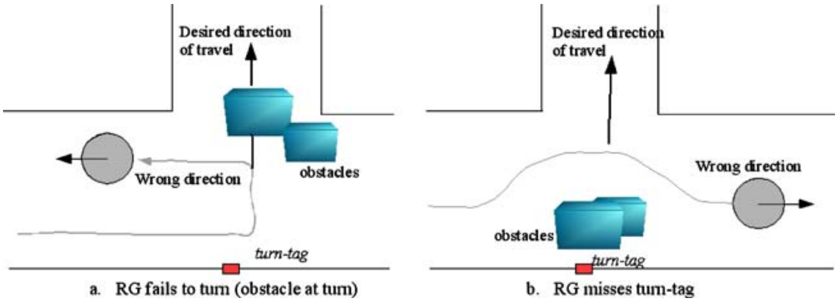


Fig. 8 Two situations leading to a loss



A loss occurs when too many invalid tags are detected. In general, RG-I always reaches B from A if the following assumptions are true: (1) the robot's batteries have sufficient power (above 8 volts); (2) there is an actual path from A to B in the current state of the world; and (3) the *critical* tags on the path from A to B are not incapacitated. By critical tags we mean the start tag, the destination tag, and the turn tags. If either the second or third assumption does not hold and the robot is lost, the loss is irrecoverable. To be more exact, the robot will keep trying to recover from the loss until its power drops down to 8 volts, which will cause a complete shutdown.

5. Pilot experiments

Our pilot experiments focused on robot-assisted navigation and human-guide interaction. The robot-assisted navigation experiments evaluated the ability of visually impaired individuals to use RG-I to navigate unfamiliar environments as well as the ability of the robot to navigate on its own. The human-robot interaction experiments investigated how visually impaired individuals can best interact with robotic guides.

5.1. Robot-assisted navigation

We deployed our system for a total of approximately seventy hours in two indoor environments: the Assistive Technology Laboratory (ATL) of the USU Center for Persons with Disabilities and the USU CS Department. The ATL occupies part of a floor in a building on the USU North Campus. The area occupied by the ATL is approximately 4,270 square meters and contains 6 laboratories, two bathrooms, two staircases, and an elevator. The CS Department occupies an entire floor in a multi-floor building. The floor's area is 6,590 square meters. The floor contains 23 offices, 7 laboratories, a conference room, a student lounge, a tutor room, two elevators, several bathrooms, and two staircases.

Forty RFID tags were deployed at the ATL and one hundred tags were deployed at the CS Department. It took one person 20 minutes to deploy the tags and about 10 minutes to remove them at the ATL. The same measurements at the CS Department were 30 and 20 minutes, respectively. The tags, which were placed on small pieces of cardboard to insulate them from the walls, were attached to the walls with regular masking tape. The creation of the connectivity graphs took one hour at the ATL and about two and a half hours at the CS Department. One member of our research team first walked around the areas with a laptop and recorded tag-destination associations and then associated behavior scripts with tags.

RG-I was first repeatedly tested in the ATL, the smaller of the two environments, and then deployed for pilot experiments at the USU CS Department. We ran three sets of navigation experiments. The first and third sets did not involve visually impaired participants. The second set did. In the first set of experiments, we had RG-I navigate three

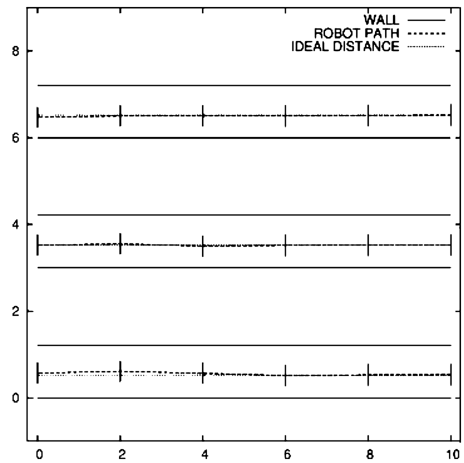


Fig. 9 Path deviations in narrow hallways

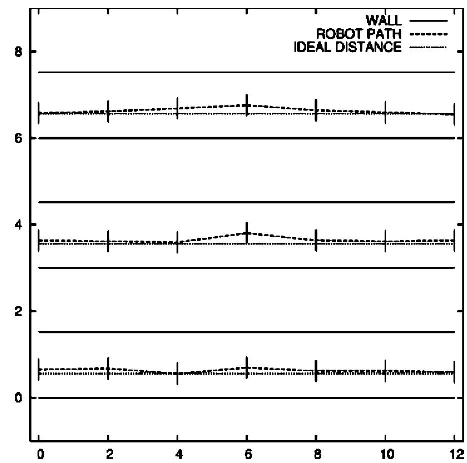


Fig. 10 Path deviations in medium hallways

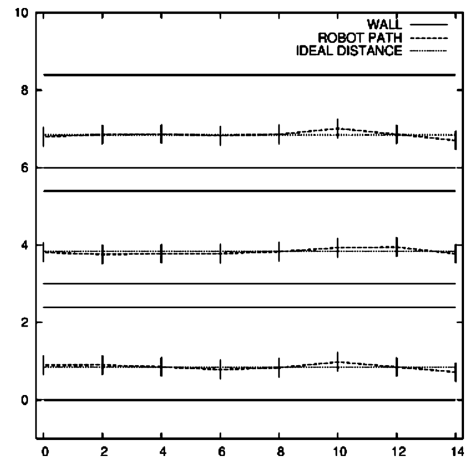


Fig. 11 Path deviations in wide hallways

types of hallways in the CS Department: narrow (1.5 meters), medium (2.0 meters) and wide (4.0 meters), and we evaluated its navigation in terms of two variables: path deviations and abrupt speed changes. We also observed how well the robot's RFID reader detected the tags.

To estimate path deviations, in each experiment we first computed the ideal distance that the robot has to maintain from the right wall in a certain width type of hallway (narrow, medium, and wide). As shown in Fig. 12, the ideal distance was computed by running the robot in a hallway of the type being tested with all doors closed and no obstacles en route. RFID tags were placed along the right wall of every route every two meters to help with interpolation and graphing. During each run, the distance read by the laser range finder between the robot and the right wall was recorded every 50 milliseconds. The ideal distance was computed as the average of the distances taken during the run. Once the ideal distances were known, we ran the robot three times in each type of hallway. The hallways in which the robot ran were different from the hallways in which the ideal distances were computed. Obstacles, e.g., humans walking by and open doors, were allowed during the test runs. The average of all the readings for each set of three runs, gives the average distance the robot maintains from the right wall in a particular type of hallway.

Figures 9–11 give the distance graphs of the three runs in each hallway type. The vertical bars in each graph represent the robot's width. As can be seen from Fig. 9, there is almost no deviation from the ideal distance in narrow hallways. Nor is there any oscillation. Figures 10 and 11 show some insignificant deviations from the ideal distance. The deviations were caused by people walking by and by open doors. However, there is no oscillation, i.e., sharp movements in different directions. In both environments, we observed several tag detection failures, particularly near or on metallic door frames. However, after we insulated the tags with thicker pieces of cardboard, the tag detection failures stopped.

Figures 13–15 give the velocity graphs for each hallway type (x-axis is time in seconds, y-axis is velocity in mm/sec). The graphs show that the narrow hallways cause short abrupt changes in velocity. In narrow hallways even a slight disorientation, e.g., 3 degrees, in the robot causes changes in velocity because less empty space is detected in the grid. In medium and wide hallways, the velocity is generally smooth. However, several speed changes occur when the robot passes or navigates through doorways or avoids obstacles.

The mean and standard deviation numbers for the hallway experiments were as follows: in wide hallways, $\mu = 708.94$, $\sigma = 133.32$; in medium hallways, $\mu = 689.19$, $\sigma = 142.32$; in narrow hallways, $\mu = 670.43$, $\sigma = 166.31$. It should be noted that the means are influenced by the fact that the robot always started at 0 velocity. Thus, since the mean, as a statistical measure, is influenced by outliers, these means may be slightly skewed.

The second set of pilot experiments involved five visually impaired participants, one participant at a time, over a period

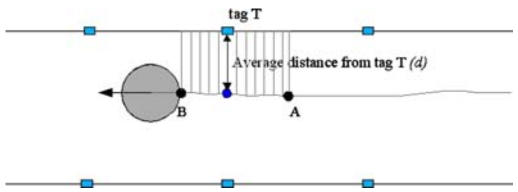


Fig. 12 Computing an ideal distance in hallways

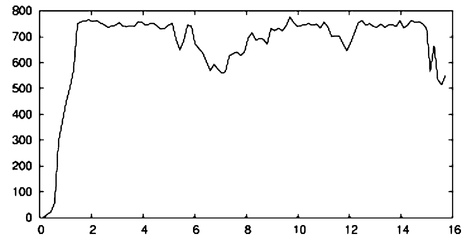


Fig. 13 Velocity changes in narrow hallways

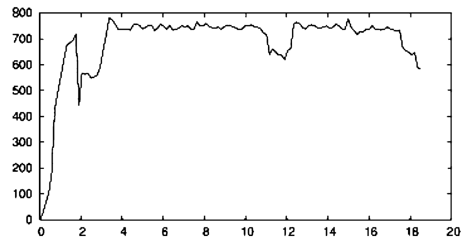


Fig. 14 Velocity changes in medium hallways

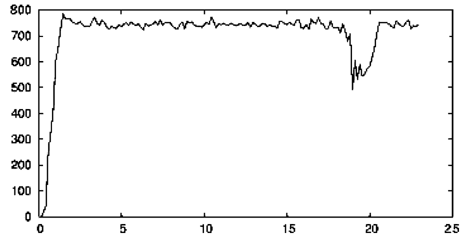


Fig. 15 Velocity changes in wide hallways

of two months. Three participants were completely blind and two participants could perceive only light. The participants had no speech impediments, hearing problems, or cognitive disabilities. Two participants were dog users and the other three used white canes. The participants were asked to use RG-I to navigate to three distinct locations (an office, a lounge, and a bathroom) at the USU CS Department. All participants were new to the environment and had to navigate approximately 40 meters to get to all destinations. Thus, in the experiments with visually impaired participants, the robot navigated approximately 200 meters. All participants had to use a wireless wearable microphone to interact with the robot: at the beginning of a run, each participant would speak the destination he or she wanted to reach. All participants reached their destinations. In their exit interviews, all participants said they liked the fact that they did not have to give up their white canes and/or guide dogs to use RG-I. Most complaints were about the human-robot interaction aspects of the system. For example, all of them had problems with the speech recognition system and often had to repeat destinations several times before the robot understood them (Kulyukin et al., 2004; Sute, 2004).

Another problem with speech recognition occurs when the person guided by RG-I stops and engages in conversation with someone. Since speech recognition runs continuously, some phrases said by the person during a conversation may be erroneously recognized as route directives, thereby causing the robot to start moving. For example, once RG-I erroneously recognized a directive and started pulling its user away from his interlocutor until the user's stop command pacified it. In another situation, RG-I managed to run a few meters away from its user, because the user hung the leash on the PVC pole when he stopped to talk to a friend in a hallway. Thus, after saying "Stop," the user had to grope his way along a wall to the robot that was standing a few meters away.

In the third navigation experiment, RG-I was made to patrol the entire area of the USU CS Department on its own. This experiment focused on recoverable and irrecoverable losses on two types of routes: (1) simple structural, simple agent-based and (2) complex structural, simple agent-based. The first type was operationally defined as routes having 0, 1, or 2 turns, of less than 40 meters in length with no or few people or obstacles. The second type was operationalized as routes with more than 2 turns, of more than 40 meters in length, with no or few people or obstacles. A total of 18 routes (9 routes of the first type and 9 routes of the second type) were set up. The robot continuously navigated these routes until its battery voltage reached 8 volts. The robot navigated a total of 6 kilometers for 4 hours and had 5 recoverable and 0 irrecoverable losses. All losses occurred on complex structural and simple agent-based routes. During several separate trial runs with visually impaired participants on different types of routes, the robotic guide suffered several recoverable losses. While we did not collect statistics on the participants' reactions to recoverable losses, we observed that the participants did not complain. Two participants said that they would not have known that the robot had to take a different route if the robot had not announced it to them.

5.2. Human-robot interaction

After we tested speech-based interaction in navigation experiments and received negative feedback from the participants, we decided to evaluate the feasibility of speech more systematically. Therefore, our first human-guide experiment tested the feasibility of using speech as a means of input for humans to communicate with the robot. Each participant was asked to speak approximately sixty phrases while wearing a headset that consisted of a microphone and one headphone. The phrase list was a list of standard phrases that a person might say to a robotic guide in an unfamiliar environment, e.g., ”go to the bathroom,” ”where am I?” etc. Each phrase was encoded as a context-free command and control grammar rule in SAPI’s XML-based grammar formalism. Each participant was positioned in front of a computer running SAPI. The test program was written to use SAPI’s text-to-speech engine to read the phrases to the participant one by one, wait for the participant to repeat a phrase, and record a recognition result (speech recognized vs. speech not recognized) in a database.

The speech feasibility experiment was repeated in two environments: noise-free and noisy. The noise-free environment did not have any ambient sounds other than the usual sounds of a typical office. To simulate a noisy environment, a long audio file of a busy bus station was played on another computer in the office very close to where each participant was sitting. All five participants were native English speakers and did not train SAPI’s speech recognition engine on sample texts.

We found that the average percentage of phrases recognized by the system in the noise-free environment was 38%, while the average percentage of recognized phrases in the noisy environment was 40.2%. Although the level of ambient noise in the environment did not seem to affect the system’s speech recognition, in both environments fewer than 50% of phrases were correctly recognized. Even worse, some non-phrases were incorrectly recognized as phrases. For example, when one participant made two throat clearing sounds, the system recognized the sound sequence as the phrase ”men’s room.”

The statistics were far better for the participants understanding phrases spoken by the computer. The average percentage of speech understood in the noise-free environment was 83.3%, while the average percentage of phrases understood in the noisy environment was 93.5%. Clearly, in the second trial (the noisy environment), the participants were more used to SAPI’s speech recognition and synthesis patterns. These results suggest that speech appears to be a better output medium than input (Sute, 2004).

Audio perception experiments were conducted with all five participants to test whether they preferred speech to audio icons, e.g., a sound of water bubbles, to signify different objects and events in the environment and how well participants remembered their audio icon selections. A simple GUI-based tool was built that allows visually impaired users to create their own audio associations. The tool was used to associate events and objects, e.g., water cooler to the right, approaching left turn, etc., with three audio messages: one speech message and two audio icons. A small number of objects was chosen to eliminate steep learning curves. All in all, there were seven different objects, e.g., elevator, vending machine, bathroom, office, water cooler, left turn, and right turn.

Each object was associated with two different events: *at* and *approaching*. For example, one can be at the elevator or approaching the elevator. The audio icons available for each event were played to each participant at selection time. The following statistics were gathered:

1. Percentage of accurately recognized icons;
2. Percentage of objects/events associated with speech;
3. Percentage of objects/events associated with audio icons;
4. Percentage of objects/events associated with both.

The averages for these experiments were:

1. Percentage of accurately recognized icons − 93.3%;
2. Percentage of objects/events associated with speech −55.8%;
3. Percentage of objects/events associated with icons −32.6%;
4. Percentage of objects/events associated with both −11.4%.

The analysis of the audio perception experiments showed that two participants were choosing audio preferences essentially at random, while the other three tended to follow a pattern: they chose speech messages for *at* events and audio icons for *approaching* events or vice versa. The experiments also showed that the participants tended to go either with speech or with audio icons, but rarely with both. The experiments did not give a clear answer as to whether visually impaired individuals prefer to be notified of objects/events via speech or audio icons. It is important to keep in mind, however, that our objective was to collect preliminary descriptive statistics on the perception of audio cues in robot-assisted navigation. No attempt was made to make statistically significant inferences. Further work is needed on a larger and more representative sample to answer this question on a statistically significant level.

6. Conclusions

From our experiences with RG-I, we can make the following preliminary observations.

First, orientation-free RFID-based navigation guarantees reachability at the expense of optimality (Kulyukin et al., 2004). If the path to the destination is not blocked and all critical tags are in place, the robot reaches the destination. The obvious tradeoff is the optimality of the path, because the actual path taken by the robot may be suboptimal in terms of time and distance due to a recoverable loss.

Second, the instrumentation of the environments with RFID sensors is reasonably fast and requires only commercial off the shelf (COTS) hardware and software components. RFID tags are inexpensive (15 USD a piece), reliable, and easy to maintain, because they do not require external power supplies. RFID tag reading failures are rare, and can be recovered from as long as a large number of tags are placed in the environment. The placement of RFID tags in the environment does not seem to disrupt any indigenous activities. People who work in the environment do not seem to mind the tags due to their small size.

Third, an alternative technique for recognizing such landmarks as left turns, right turns, and T- and X-intersections will make the navigation behavior more robust even when critcal tags are not detected. We are investigating several landmark recognition techniques that work on laser range finder signatures. Another problem that we plan to address in the future is the detection of irrecoverable losses. As of now, RG-I cannot detect when the loss is irrecoverable.

Fourth, the robot is able to maintain a moderate walking speed during most of the route, except at turns and during obstacle avoidance. The robot's motion is relatively smooth, without sideways jerks or abrupt speed changes. However, obstacles that block a critical tag may cause the robot to miss the tag due to obstacle avoidance and fail to trigger an appropriate behavior. In addition, at intersections, RG-I can select a wrong open space due to obstacles blocking the correct path. Adding other landmark recognition techniques is likely to improve the robot's navigation at intersections.

Fifth, at this point in time speech, when recognized by Mircrosoft's SAPI 5.1 with no user training, does not appear to be a viable input mode. As we indicated elsewhere (Kulyukin, 2004; Sute, 2004), it is unlikely that speech recognition problems can be solved on the software level until there is a substantial improvement in the state-of-the-art speech recognition. Our pilot experiments suggest that speech appears to be a viable output mode. We believe that for the near future wearable hardware solutions may offer reliable input modes. We are currently exploring human-robot interaction through a wearable keypad. The obvious advantage is that keypad-based interaction eliminates the input ambiguity problems of speech recognition. Additional experiments with human participants are needed to determine the feasibility of various wearable hardware devices for human-robot interaction.

The previous conclusion is not to be construed as an argument that speech-based HRI is not an important venue of research. It is. However, it is important not to confuse interaction itself with a specific mode of interaction. Speech is just one mode of interaction. Typed text, eye gaze, sipping and puffing, gesture and touch are also valid interaction modes. As assistive technology researchers, we are interested, first and foremost, in effective and safe communication between a disabled person and an assistive device. Consequently, to the extent that speech introduces ambiguity, it may not be appropriate as an interaction mode in some assistive robots.

Finally, the SSH construction is done manually. The obvious question is can it be completely or partially automated? We are currently investigating a tool that would allow one to generate a tag connectivity graph and associate tags with behavior scripts through drag and drop GUIs. Another possibility that we are contemplating is equipping RG-I with an orientation sensor and manually driving the robot with a joystick on the previously chosen routes in an environment with deployed RFID tags. As the robot is driven through the environment, it senses the RFID tags and turns and associates the detected tags with behavior scripts. In effect, the SSH is first constructed as the robot is driven through the environment and is subsequently edited by a knowledge engineer.

Acknowledgements The first author would like to acknowledge that this research has been supported, in part, through the NSF CAREER Grant (IIS-0346880), two Community University Research Initiative (CURI) grants from the State of Utah, and a New Faculty Research grant from Utah State University. The authors would like to thank the reviewers of this article for their helpful and instructive comments.

References

1. Addlesee, M. Curwen, R., Hodges S., Newman J., Steggles P., and Ward A. 2001. Implementing a sentient computing system. *IEEE Computer*, August:2–8.
2. Agre, P. 1988. The dynamic structure of everday life. Ph.D. Thesis. MIT Artificial Intelligence Laboratory.
3. Benjamin, J. M., Ali, N. A., and Schepis, A. F. 1973. A Laser cane for the blind. In *Proceedings of San Diego Medical Symposium*.
4. Bissit, D. and Heyes, A. 1980. An application of biofeedback in the rehabilitation of the blind. *Applied Ergonomics* 11(1):31–33.
5. Borenstein, J. and Koren, Y. 1989. Real-time obstacle avoidance for fast mobile robots. *IEEE Transactions on Systems, Man, and Cybernetics* 19:1179–1187.
6. Borenstein, J. and Ulrich, I. 1994. The guidecane—A computer. In *Proceedings of the IEEE International Conference on Robotics and Automation*, San Diego, CA.
7. Burgard, W., Cremers, A., Fox, D., Hähnel, D., Lakemeyer, G., Schulz, D., Steiner, W., and Thrun, S. 1999. Experiences with an interactive museum tour-guide robot, *Artificial Intelligence*, 114:3–55.
8. Fong, T. and Thorpe, C. 2001. Vehicle teleoperation interfaces. *Autonomous Robots*, 11(2):9–18.
9. Gharpure, C. 2004. Orientation Free RFID-Based Navigation in a Robotic Guide for the Visually Impaired, Masters Thesis. Department of Computer Science, Utah State University.
10. Horswill, I. 1993. Polly: A vision-based artificial agent. In *Proceedings of the Conference of the American Association for Artificial Intelligence (AAAI-1993)*, Washington, DC.
11. Kantor, G. and Singh, S. 2002. Preliminary results in range-only localization and mapping. *Proceedings of the IEEE Conference on Robotics and Automation (ICRA-2002)*, Washington, DC.
12. Khatib, O. 1985. Real-time obstacle avoidance for manipulators and mobile robots. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA-1985)*, St. Louis, Missouri.

13. Koren, Y. and Borenstein, J. 1991. Potential field methods and their inherent limitations for mobile robot navigation. In *Proceedings of the IEEE Conference on Robotics and Automation (ICRA-1991)*, Sacramento, CA.
14. Kulyukin, V. and Blair, M. 2003. Distributed tracking and guidance in indoor environments. In *Proceedings of the Rehabilitation Engineering and Assistive Technology Society of North America (RESNA-2003)*, Atlanta, GA, avail. on CD-ROM.
15. Kulyukin, V., Sute, P., Gharpure, C., and Pavithran, S. 2004. Perception of audio cues in robot-assisted navigation. In *Proceedings of the Rehabilitation Engineering and Assistive Technology Society of North America (RESNA-2004)*, Orlando, FL, avail. on CD-ROM.
16. Kulyukin, V., Gharpure, C., and De Graw, N. 2004. Human-Robot Interaction in a Robotic Guide for the Visually Impaired. In *Proceedings of the AAAI Spring Symposium on Interaction between Humans and Autonomous Systems over Extended Operation*, Stanford, CA.
17. Kulyukin, V. 2004. Human-Robot Interaction through Gesture-Free Spoken Dialogue. *Autonomous Robots* 16(3):239–257.
18. Kulyukin, V., Gharpure, C., Nicholson, J., and Pavithran, S. 2004. RFID in robot-assisted indoor navigation for the visually impaired. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Systems and Robots (IROS-2004)*, Sendai, Japan.
19. Kulyukin, V., Gharpure, C., and Nicholson, J. 2005. RoboCart: Toward robot-assisted navigation of grocery stores by the visually impaired. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Systems and Robots (IROS-2005)*, Edmonton, Alberta, Canada.
20. Kupiers, B. 2000. The spatial semantic hierarchy. *Artificial Intelligence* 119:191–233.
21. LaPlante, M. and Carlson, D. 2000. *Disability in the United States: Prevalence and Causes*. National Institute of Disability and Rehabilitation Research, U.S. Department of Education: Washington, D.C.
22. Montemerlo, M., Pineau, J., Roy, N., Thrun, S., and Verma, V. 2002. Experiences with a mobile robotic guide for the elderly. In *Proceedings of the Conference of the American Association for Articial Intelligence (AAAI-2002)*, Edmonton, AB.
23. Mori, H. and Kotani, S. 1998. Robotic travel aid for the blind: HARUNOBU-6. *Second European Conference on Disability, Virtual Reality, and Assistive Technology*, Sövde, Sweden.
24. http://www.opencyc.org. 2003. *The OpenCyc Project: Formalized Common Knowledge*. Cycorp, Inc..
25. Pfaffenberger, C., Scott, J. P., Fuller, J., Ginsburg, B. E., and Bielfelt, S. W. 1976. *Guide Dogs for the Blind: Their Selection, Development, and Training*. Elsevier Scientific Publishing: Amsterdam, Holland.
26. Ross, D. 2001. Implementing assistive technology on wearable computers. *IEEE Intelligent Systems* May:2–8.
27. Ross, D. and Blasch, B.B. 2002. Development of a wearable computer orientation system. *IEEE Personal and Ubiquitous Computing*, 6:49–63.
28. Simmons, R. 1996. The Curvature-Velocity method for local obstacle avoidance. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA-1996)*, Minneapolis, USA.
29. Shoval, S. Borenstein, J., and Koren, Y. 1994. In Mobile robot obstacle avoidance in a computerized travel for the blind. In *Proceedings of the IEEE International Conference on Robotics and Automation*, San Diego, CA.
30. Thrun, S., Bennewitz, M., Burgard, W., Cremers, A. B., Dellaert, F., Fox, D., Hähnel, D., Rosenberg, C., Roby, N., Schutle, J., and Schultz, D. 1999. Minerva: A second generation mobile tour-guide robot. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA-1999)*, Antwerp, Belgium.
31. Tinbergen, N. 1976. *Animal in its World: Laboratory Experiments and General Papers*. Harvard University Press.
32. Tsukiyama, T. 2003. Navigation system for the mobile robots using RFID tags. In *Proceedings of the IEEE Conference on Advanced Robotics*, Coimbra, Portugal.
33. Hähnel, D., Burgard, W., Fox, D., Fishkin, K., and Philipose, M. 2003. *Mapping and Localization with RFID Technology*. Intel Research Institute, Seattle, WA.
34. Sute, P. 2004. Perception of Audio Cues in Robot-Assisted Navigation for the Visually Impaired, Masters Report. Department of Computer Science, Utah State University.