Mobile Manipulators for Assisted Living in Residential Settings

Patrick Deegan, Roderic Grupen, Allen Hanson, Emily Horrell, Shichao Ou,
Edward Riseman, Shiraj Sen, Bryan Thibodeau, Adam Williams, Dan Xie

Department of Computer Science
University of Massachusetts Amherst
grupen@cs.umass.edu

Abstract: We describe a methodology for creating new technologies for assisted living in residential environments. The number of eldercare clients is expected to grow dramatically over the next decade as the baby boom generation approaches 65 years of age. The UMass/Smith ASSIST framework aims to alleviate the strain on centralized medical providers and community services as their clientele grow, reduce the delays in service, support independent living, and therefore, improve the quality of life for the up-coming elder population. We propose a closed loop methodology wherein innovative technical systems are field tested in assisted care facilities and analyzed by social scientists to create and refine residential systems for independent living. Our goal is to create technology that is embraced by clients, supports efficient delivery of support services, and facilitates social interactions with family and friends. We introduce a series of technologies that are currently under evaluation based on a distributed sensor network and a unique mobile manipulator (MM) concept. The mobile manipulator provides client services and serves as an embodied interface for remote service providers. As a result, a wide range of cost-effective eldercare applications can be devised, several of which are introduced in this paper. We illustrate tools for social interfaces, interfaces for community service and medical providers, and the capacity for autonomous assistance in the activities of daily living. These projects and others are being considered for field testing in the next cycle of ASSIST technology development.

Keywords: assistive robotics, distributed sensor-effector network, mobile manipulator

1 Introduction

"The United States is about to experience the greatest demographic change in its history. Most of this change will occur over the next 30 years, as 77 million baby boomers cease to work and pay payroll taxes and instead start to retire and collect benefits." (Kotlikoff, Fehr, & Jokisch 2005)

The problems arising from this change in demographics include spiraling health care costs and shortages of trained nurses and doctors (Zarit & Zarit. 1998). The percentage of elders in the population will increase dramatically when the first segment of the baby boom cohort becomes 65 in 2011 (Hobbs & Stoops 2002). Institutional support for this population will become impossible and many of these elders will want to remain in their homes, but with aging comes higher rates of functional and cognitive deficits (Gist & Hetzel 2004) that, in many cases, result in limitations in at least one activity of daily living (Kassner 2006).

It is likely that technologies for residential and assisted living settings can help to relieve the inevitable stress on the medical infrastructure and to extend the period of time that elders can live independently. First and foremost, these technologies include assistance in the activities of daily living (ADL)—technologies that enhance safety and security, assist in daily medical compliance, and help with client calendars and daily chores such as household cleaning and grocery shopping. Second, elders are susceptible to isolation as they become less mobile (Pin *et al.* 2005) (Michael *et al.* 1999). Devices that facilitate communication and social relationships between peers, families, and the surrounding community can help these clients remain connected socially. Third, a diminished capacity to travel independently means that more of this population must receive regular medical checkups in their homes by health care practicioners rather than in a centralized facility. The dearth of trained physicians and nurses can be compensated by technologies that enable "virtual" house calls. Technologies in the home that create an appropriate interface between the medical industry and the elderly client can help to make efficient use of the medical infrastructure and improve the frequency of care and oversight.

In this paper, we review developments in the research community that have considered these challenges and compare these efforts with our own. The ASSIST framework is introduced to provide an experimental environment for rapid development and testing of eldercare applications (including robotic assistants) in realistic scenarios. Preliminary findings from ASSIST have already identified and suggested technologies that are both useful and practical, and most importantly are *desired* by the potential clientele. Based on these results, several areas have been identified as particularly ripe for future development. The array of new technology being developed within our program and prepared for field deployment is presented in Sections 3 and 4. The base technologies for these applications include an active sensor-effector network and a unique, new mobile manipulator device that can function as a personal assistant in human residential environments. Built on these are several applications designed to assist in the activities of daily living, facilitate social interactions with family and the surrounding community, and enhance interactions with medical providers. Finally, we will summarize the conclusions of the study and discuss some opportunities for future work.

2 Technology Development for Assisted Living

2.1 Related Work

Our effort is aimed at creating new concepts and applications for assistive robotic systems that respond to the needs of the client, that unburden an increasingly overtaxed health care infrastructure, that allow elderly clients to live independently for a longer period of time, that increase the frequency and quality of interactions with family and community, and that facilitate the efficient delivery of medical services.

Figure 1: Pearl, the CMU Nursebot, is able remind elders of their activities and lead them to of activity if necessary

There are a wide range of goals driving the development of robots for health-care, physical and cognitive assistance, and aids for daily living. Several sophisticated humanoid robots (e.g. Asimo and Qrio) have been demonstrated (Sakagami *et al.* 2002; Tanaka & Suzuki 2004) that may one day be applicable in service robotics applications. However, they do not yet possess the right combination of price point, functionality, and the capacity for delivering work that we believe is necessary. Other system concepts have focused much more exclusively on target functionalities. For example, the CMU Nursebot (Pollack 2002b), Pearl, was designed to interact with elders in assisted living facilities and nursing homes. The robot was deployed in field tests in the Longwood Retirement Community in Oakmont, PA. It reminded elders of planned activities on their calendars and would lead them to the activity if necessary. A robotic walker (Morris *et al.* 2003) was designed for the same project that offered mobility assistance while preserving the client's fine motor control using a haptic device. The research conducted on the Nursebot focused on autonomous navigation (Pollack 2002b), activity planning (Pollack 2002a), cognitive prosthetics, and human-robot interfaces. Nursebot had an expressive head and a touch screen for client interactions. However, Nursebot is not able to interact with the environment manually and so it could not effect work on behalf of the client. Moreover, it could not function as the local embodiment of a remote service provider (family member, doctor or therapist). Our goal is to focus on just these kinds of services.

InTouch Health (http://www.intouchhealth.com/) markets the RP-7 robots to support efficient interaction between doctors and possible stroke patients in an emergency room setting. They stand about 5 feet tall, have a touch screen for a head as shown in Figure 2, and allow a remote doctor to drive the vehicle using a teleoperator interface. Mobility significantly enhances more traditional teleconferencing technologies in this application and RP-7 promises to decrease the average amount of time that a stroke patient must wait before being treated—the most significant factor affecting outcome.

A similar robot named Pebbles serves a symmetrical role—rather than acting as a surrogate for the service provider, Pebbles embodies the client and takes them (virtually) to places they could not otherwise go (Williams *et al.* 1998).

Robot devices are often considered for theraputic/rehabilitative functions. The vast majority of robotic rehabilitation work takes a hands-on approach where, for example, the robot aids the movement of a patient's limb during theraputic exercises (Leifer 1981; Burgar 2000). Mataric et al., (Tapus, Tapus, & Mataric 2007; Tapus & Mataric 2006; Tapus, Mataric, & Scassellati 2007) proposed a "hands-off" approach where the therapist robot assists the exercise of the patient via social interactions and

Figure 2: Neurosurgeon Richard Fesler is shown on the monitor of the RP-7 remote presence robotic system at St. Joseph Mercy Hospital in Pontiac, Michigan. Doctors use a laptop and the Net to connect to the robot.

Figure 3: Pebbles is a mobile teleconferencing application that attends school for disabled children—allowing them to maintain relationships with their peers. Pebbles includes a rudimentary manipulator for interacting with the teacher in the classroom setting.

encouragement. Several companion robots in the form of small animals such as a seal, a cat (Heerink *et al.* 2006), or a huggable teddy bear (Stiehl 2006) have been designed and used to demonstrate the therapeutic value of "pets" to reduce the client's stress and depression.

The applications surveyed in this section have successfully demonstrated different aspects of assistive robots in therapeutic settings. They assist with physical therapy, provide social contexts for client interactions, and contribute to the efficient use of sparse human resources. However, each application is highly specialized functionally and employs special purpose interfaces. Moreover, there has not been a personal robot that can assist in generic manual tasks, in which motor function contributes to effective user/client interfaces, or that may act as the local embodiment of remote users with different goals, including family, community services, and medical services.

2.2 The ASSIST Framework

To spur the development of relevant technology, we argue that a closed-loop process in which innovative research is combined with assessment in the field is indispensible. Figure 4 illustrates the closed loop relationship between development, implementation, and testing as adopted in this project. Our goal is to engage computer science researchers, social scientists and gerontologists to evaluate how technology is adopted and its impact on well being and healthcare delivery for the elderly. The schematic diagram portrays a framework wherein: (1) scientists and engineers prototype technology, (2) candidate technology is field tested in residential care settings in the context of clients and remote service providers, and (3) social scientists assess the net effectiveness of technology and user satisfaction *in situ*, which in turn, drives new development. The goal is to construct a methodology for developing assistive robotic technologies grounded in the needs of the client and the requirements of family members, social workers, medical and community service providers. This is achieved by ongoing focus groups engaging clients, their families, and professional caregivers in evaluating the effectiveness of candidate technologies.

An important aspect of assistive technology in residential environments is the realization that all installations are different—technology delivered into the residential environment must adapt to special needs, lifestyles, preferences, residential geometry and environment. Moreover, these technologies must also facilitate interaction between the client, family, community, and medical providers (where interfaces are more standardized). This general structure is shown in Figure 5.

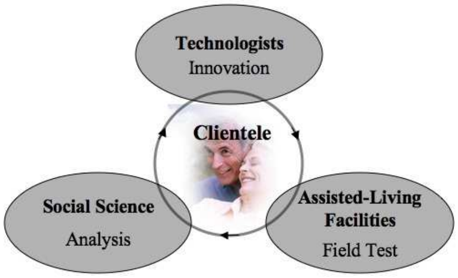


Figure 4: The closed-loop ASSIST methodology; innovation, field testing, analysis.

With the client at the center of the framework, we propose a conceptual architecture in which technological aids facilitate daily living and effect the interaction between client and remote services. An "active environment," consisting of a distributed sensor network and a novel mobile manipulator implements these client services and provides assistance for independent living. It includes core services that track client activites, implement client interfaces, and provide physical and cognitive prosthetics. The objective is to improve quality of life while simultaneously reducing the strain on resources (human and otherwise) associated with community support services. We are ensuring that the intended recipients of the technology—elder-care clients themselves—are in the design loop, and that technological solutions are effective on a case-by-case basis. We refer to this development environment as the ASSIST *sandbox*—it provides a platform for technology developers, clients, service providers, families, and payers to "play" with new concepts in assistive robotics and residential healthcare delivery.

Project ASSIST is a multi-institutional and interdisciplinary research project that aims to explore how technology can best be shaped to support the needs of elderly clients and others in assisted living settings. Members of the research team include the University of Massachusetts Computer Science Department (computer vision, robotics, distributed sensor systems, human-computer interaction), the Smith College School for Social Work (social science and geriatric social work), local elder community centers in western Massachusetts, and the Veteran's Administration (Connecticut Health Care System, West Haven campus). The project is similar in spirit and concept to the Create project (Czaja *et al.* 2002)(Czaja & Sharit 2003)(Rogers & Fisk 2004), for example, that combines computer and social sciences with the social work community. Projects such as these strive to understand the expected outcome of human-robot systems before large scale deployment of new technology.

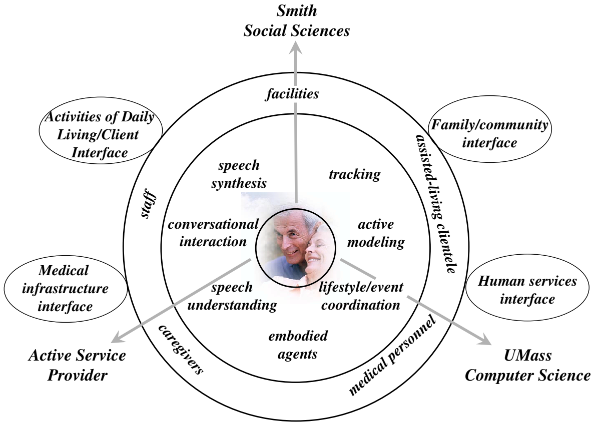


Figure 5: A multilayered architecture for creating and delivering appropriate, individualized support for independent living that facilitates client interaction with community institutions, social services, medical providers, and family.

2.3 Results from Previous ASSIST Studies

Project ASSIST has conducted four focus groups thus far, the results of which include several consistent recommendations regarding focus and delivery of first generation products. These include a fall detection system, a custom daily activity reminder system, and custom interfaces tailored specifically to elderly clientele for communicating with friends, family, and service providers. Focus groups engaged subjects that are (1) healthcare professionals, (2) elders of both genders covering a spectrum of ages, economic groups, and educational background, or (3) members of the elders' families. Opinions were solicited from the group collectively as well as from individuals. Demonstration videos of potential applications were shown first during each focus group session. After each demonstration, respondents discussed cost, functionality, interface complexity, and the special-purpose versus general-purpose character of the component technologies. An example of this last issue concerned the detection of a fall event. Several inexpensive pendant devices that are worn around the neck are available on the market for communication in the event of a fall or other medical emergency. These devices must be worn to be effective and are special purpose in the sense that they perform a single function. In contrast, our implementation (Section 3.1.2) focused on modeling activities in a distributed sensor network. This approach can be used to support a variety of other functions in addition to the fall event. There are interesting cost/benefit questions that arise based on decisions like these that can only be addressed when integrated systems are evaluated in the field for a variety of client populations. To our knowledge, this has yet to be done for the kinds of problems and application we are considering.

Our focus groups also revealed other preferences. The elder focus group participants were very enthusiastic about the use of video technology (e.g. the videophone), as they anticipate future impairment that overrides their own personal concerns about privacy. This finding is consistent with some other reports in the literature (Lachman & Andreoletti 2006). Elders also seemed less fearful about dealing with technology that they don't yet understand if they believe the technology will benefit them in some way (e.g. improve health care, support an independent lifestyle, or improve safety). They also appreciated access to the scientists and graduate students involved in the project and their willingness to answer questions and respond to concerns. This implies that our closed-loop methodology may improve both the early adoption of new technology by the target population and the technology itself. Preliminary conversations with focus group attendees in informal/open Q&A session included enthusiastic support for active agents in the form of small robots that could be deployed in the home. From these conversations, it appears important to expand services beyond the *detect-and-alert* model, to include effectors that perform work in the environment and that can interact with the user in a physical way. For example, a mobile manipulator that can take preemptive action to reduce the probability that the client will trip and fall by cleaning up clutter was viewed as an attractive option.

The constraints imposed by interfaces was a frequent issue. Elders will not and arguably should not spend a great deal of time in front of the computer, but should be encouraged to move around the house performing safe and healthy activities. Studies from Haigh (Haigh, Kiff, & Ho 2006) and Koester (Koester 2004) have shown that interactive speech interfaces can be confusing to some elders when they hear a voice in the room that is disembodied from any focal agent, suggesting that maybe a desktop interface could be advisable despite lifestyle issues. As result, we are exploring an alternative interface that involves a mobile robot serving in part as a "mobile communication kiosk" with an onboard touch screen that can follow a client around while he/she is engaged in other tasks. This would serve as an embodied interface for the distributed system, while providing flexible remote access to the range of client services offered. Again, cost and safety will determine whether this route is feasible.

3 Core ASSIST Technologies

This section focuses on descriptions of two of the more interesting of the core technologies comprising the ASSIST system. The distributed sensor array provides tracking and localization services, face detection and identification services for recognition of valid system users and other security concerns, activity modeling, and serves as the distributed "eyes" for the mobile manipulator. The second major core technology is a mobile manipulator that is both a physical prosthetic and an embodied interface. It serves as a cognitive focal point for the client and as a surrogate for remote service providers that is capable of natural language communication with anthropomorphic gestures and large scale movement. Moreover, it summarizes services in a distributed array of resources in a single voice that avoids the negative implications and disorientation that come with a disembodied voice. The ASSIST system represents a significant opportunity for facilitating interactions between the client, the distributed sensor/effector network, and remote services/family that exploits the innate ability of an embodied system to command and direct attention.

3.1 Distributed Sensor Array

Sensor and actuator networks are gaining increasing attention in cross disciplinary research in robotics and networking due to promising real-world applications that include environmental monitoring (Martinez, Hart, & Ong 2004), robotic swarms (Kumar, Rus, & Singh 2004), agricultural management (Burrell, Brooke, & Beckwith 2004), habitat monitoring (Mainwaring *et al.* 2002), homeland security and transportation (Li *et al.* 2002), and eldercare (Noury *et al.* 2000).

To install a sensor network in a residential healthcare situation, the first challenge is to design a configuration geometry for independent sensors. A good design provides maximal coverage using minimal resources. This sensor configuration must be scheduled in a manner that serves the needs of other clients in a dynamic environment. These two design specifications depart significantly from the current emphasis on system level issues, such as communication (Ganesan *et al.* 2001), power consumption, miniaturization, scalability (Martinez, Hart, & Ong 2004), and security (Eschenauer & Gligor 2002). This section describes the ASSIST research aimed at developing practical and deployable sensor networks in residential healthcare applications. We have focused on services for localizing and tracking features in indoor environments.

3.1.1 Localization Service The ability to localize and track a subject with a distributed array of cameras depends directly on the geometry of the sensors involved. Badly configured sensors can be occluded or otherwise ill-conditioned for the kinds of queries required by the applications they must support. We have constructed a system in which observations by a camera network are used to build models of the dynamics of human activities. Among other things, these models support the allocation of sensors to localize and track the motions of humans in their living space (Karuppiah *et al.* 2005). Figure 6 illustrates our approach. The left panel shows an experimental floorplan within which 5 independent sensors are positioned.

Given a fixed stereo geometry, observational utility over areas of the environment can be determined offline in terms of two properties: (1) their binocular field of view (FOV); and (2) the expected precision of triangulation. The position, orientation, and FOV of a pair of cameras determine their binocular FOV (Figure 6(b)). We represent this region for a camera pair, R_{ij}, as a uniform probability density, $\boldsymbol{\kappa}_o$, over $\boldsymbol{p}$ (which are centers of discretized grid cells) as follows:

$$\boldsymbol{\kappa}_o(\boldsymbol{p}, R_{ij}) = \begin{cases} \epsilon & \text{if } \boldsymbol{p} \text{ is within the overlap region} \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

where ϵ is $1/A$ and A is the total area of the region of overlap. The camera pair, R_{ij}, is viable as long as $\boldsymbol{\kappa}_o(\boldsymbol{p}, R_{ij}) > 0$.

We use the stereo localizability metric to evaluate the triangulation quality of different camera pairs. The localizability metric is defined to be the instantaneous uncertainty in the location of a subject. If B is the baseline between two cameras and θ_1 and θ_2 are the respective headings to the subject, the uncertainty Jacobian is given as follows:

$$J = \frac{B}{\sin^2(\theta_2 - \theta_1)} \begin{pmatrix} \cos\theta_2 \, \sin\theta_2 & -\cos\theta_1 \, \sin\theta_1 \\ \sin^2\theta_2 & -\sin^2\theta_1 \end{pmatrix} \tag{2}$$

The instantaneous localizability measure is $\boldsymbol{\kappa}_p(\boldsymbol{p}, R_{i,j}) = \sqrt{|JJ^T|}$, where $|\cdot|$, is the determinant of a matrix (Karuppiah *et al.* 2005). The lower this measure is, the higher the precision of localization. Depending on the cost associated with switching to a different pair, the current pair may continue to be used as long as the uncertainty is within acceptable limits. The metric $\boldsymbol{\kappa}_p$ for camera pair 0 and 1 is illustrated in Figure 6 (c).

The product of the FOV and localizability fields provides a quality measure based solely on the position of the subject.

$$\boldsymbol{\kappa}(\boldsymbol{p}, R_{ij}) = \boldsymbol{\kappa}_o(\boldsymbol{p}, R_{ij}) \cdot \boldsymbol{\kappa}_p(\boldsymbol{p}, R_{ij}) \tag{3}$$

Figure 6 (d) shows this composition.

The allocation of sensors to track the *motion* of a human subject can be influenced by the observed velocity of the subject as well as the position of the sensors. We use kernel density estimation to model the activity density. The probability of observing motion (or activity) at a grid location $\boldsymbol{p}$ using a resource R_{ij} is given by

$$\boldsymbol{\xi}_{ij}(\boldsymbol{p}) \propto \sum_{k=1}^{N_{ij}} K_h(\boldsymbol{p} - \boldsymbol{p}_k) \tag{4}$$

where $\boldsymbol{p}_k$ are the locations where motion was observed, N_{ij} is the total number of observations made by R_{ij} and $K_h(\cdot)$ is a suitable kernel function with a bandwidth parameter h. The overall probability of observing motion at a grid location from all resources is given by

$$\boldsymbol{\xi}(\boldsymbol{p}) = \sum_{j=1}^{N} K_h(\boldsymbol{p} - \boldsymbol{p}_j) \tag{5}$$

where $N = \sum_i N_i$ is the total number of observations made. Finally, we can compute the activity weighted static quality of a sensor pair:

$$\begin{aligned} \boldsymbol{\nu}(\boldsymbol{p}, R_{ij}) &= \Pr(\boldsymbol{p}|R_{ij}) \cdot \boldsymbol{\kappa}(\boldsymbol{p}, R_{ij}) \\ &= \frac{\boldsymbol{\xi}_{ij}(\boldsymbol{p})}{\boldsymbol{\xi}(\boldsymbol{p})} \cdot \boldsymbol{\kappa}(\boldsymbol{p}, R_{ij}). \end{aligned} \tag{6}$$

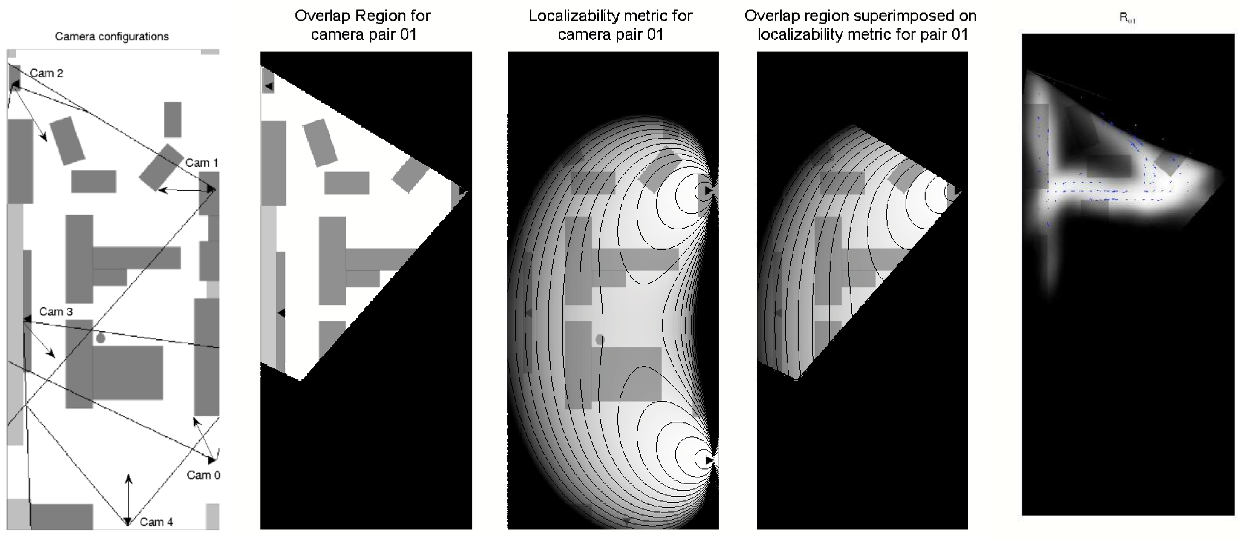
Note that, if $\boldsymbol{p}$ is in the invalid region of the cameras in R_{ij} as determined by the offline utility measure $\boldsymbol{\kappa}$, then $\boldsymbol{\nu}(\boldsymbol{p}, R_{ij})$ is zero.

The kernel function describing the quality of a sensor pair at position $\boldsymbol{p}$ provides an estimate over a neighborhood of support near $\boldsymbol{p}$. An anisotropic kernel function is chosen that is dilated in the direction of subject velocity. For each $\boldsymbol{p}$, we compute the mean $\vec{V}(\boldsymbol{p})$ over all motion vectors passing through that node in the grid. A uniform kernel K_h is scaled and aligned with $\vec{V}$ to weight the quality of neighboring positions in these directions more heavily. In this way, estimates of quality used for sensor allocation respect exemplary *activities* in the environment. Figure 6 (e) illustrates the activity weighted quality estimate for camera pair R_{01} .

Thus, for every observed subject position and velocity, we can estimate the camera pair with maximum current *and* future utility:

$$R^*(\boldsymbol{p}, \vec{V}) = \arg\max_{R_{ij}} \boldsymbol{\nu}(\boldsymbol{p}, R_{ij}) \tag{7}$$

Residential health care environments and clientele can vary significantly. Sensor networks applied to these applications must be able to accommodate this variation and maximize the quality of information gathered. Quality depends in part on stationary features of the camera and environmental geometry. It also depends on client activities over extended periods of time. We have developed a technique for modeling routinely occurring events and scoring sensor allocation strategies based on the “cost” of tracking in different regions of the room.



(a) (b) (c) (d) (e)

Figure 6: Abstractions of floorspace for localization and tracking: (a) the floorplan and the placement of cameras; (b) the overlapping binocular FOV for camera pair 01; (c) the binocular localizability metric; (d) the FOV *masked* localizability metric; (e) the anisotropic weighted activity density that reflects observed velocity.

3.1.2 Activity Modelling *Activities* are characteristic temporal patterns in observations extracted from the caregiver environment through the distributed sensor network. They can include visual, auditory, and haptic feedback as well as events generated by feedback control processes. All such events can be treated uniformly using robust supervised learning techniques involving Hidden Markov Models(HMMs). HMMs are used for modeling generative sequences which can be characterized by an underlying process generating an observable sequence. In order to detect a particular class of activity, e.g. walking, sitting, falling, etc., time-series training data must be collected for each class. The EM algorithm (Dempster, Laird, & Rubin 1977) is then used for finding the maximum-likelihood estimate of the parameters (λ) of the model, given a set of observable feature vectors, O.

$$\lambda^* = argmax_\lambda p(O|\lambda) \tag{8}$$

The HMM-based maximum likelihood procedure can also be used to recognize multiple (> 2) classes of activities, e.g. walking vs. sitting vs. lying down vs. falling. The effectiveness of this technique depends largely on how similar one class of activity is to the next.

In addition to modeling specific classes of activity, an HMM can also model the usual patterns of activity in a particular environment. This is done simply by taking a representative sample of normal activities and using them to train a single HMM. This HMM can be used to determine the likelihood of any activity occurring in a given location under normal circumstances. In the following section, this technique will be employed to detect when a potentially hazardous obstacle has been placed in a commonly trafficked area. Details of the method can be found in (Williams *et al.* 2006).

3.2 Embodied Mobile Manipulator

An important feature of our distributed sensor/effector network is an integrated mobile manipulator that can move about the residential environment and perform manual tasks. This platform carries a web cam, a microphone, a speech synthesizer, and a touch sensitive LCD display to act as a mobile and teleoperable interface. As such, it becomes a cognitive focal point for the client and a surrogate for remote service providers capable of gesture, motion, and verbal communication. Figure 7 illustrates most of these features in the uBot-4 mobile manipulator.

Even though the uBot-4 was designed primarily for research, design specifications make it a useful protype mobile manipulator for many residential assistive applications.

Mobility - The uBot-4 is a small and lightweight dynamically balancing mobile manipulator. Mobility is provided by two wheels in a differential drive configuration. While legged robots may be capable of addressing more types of terrain in principle, legs introduce much greater cost and complexity. A wheeled

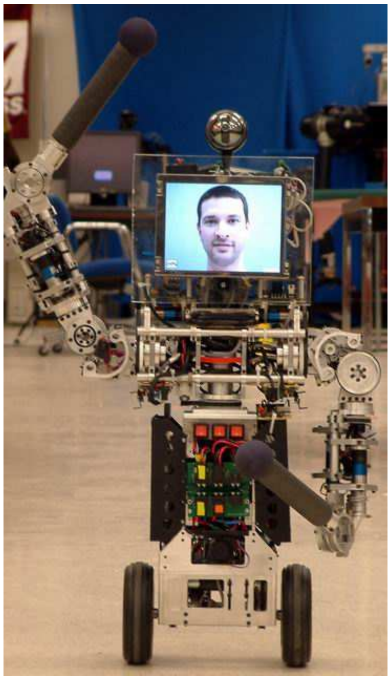


Figure 7: The UMass uBot-4 is designed to be a surrogate for an array of remote service providers and to be capable of autonomous client services.

robot can traverse almost any terrain that is accessible to an ambulatory client with a moderate degree of disability. It will use elevators instead of stairs, but has no problem with different types of flooring and doorway thresholds.

Manipulation - The uBot-4 has two four degree of freedom arms and a rotating trunk. Each arm is roughly 0.5 meters in length and the bimanual workspace contains a significant portion of the ground plane. The arms are strong enough to brace when the platform is destabilized and to do a pushup to return a vertical posture from the prone position if it should fall down.

Safety - The uBot-4's small size and relatively low mass decreases the energy released by unplanned "bumps" and reduces the probability of damage to itself, the client, or the environment. These tradeoffs mean that it will not initially be able to reach the kitchen cabinets, however, safety concerns for this prototype were considered paramount. One implication of the choice to realize a dynamically balancing robot is that the longitudinal impedance of the platform is governed by the impedance of an inverted pendulum. Careful planning, in the form of postural control, can turn the compensated inverted pendulum from a low impedance device into a forceful and relatively stiff device.

Cost - These kind of devices must be tested and proven successful before they will enjoy the economies generated by manufacturing volume. For research prototypes, it is important that they do not require a great deal of capital investment to get answers. More economical platforms will support more researchers to study the important issues that must be resolved. Moreover, reducing the cost lowers the barriers to early adopters from which a great deal of useful information can be gathered. The uBot-4 prototype costs roughly $15,000 in parts.

Performance - The robot should perform work on a time scale comparable to human performance. The uBot-4 can raise its outstretched arm with a load of about 1 kg, it can translate on a flat floor at approximately 3 mph, and has a hierarchical control architecture so that it can respond quickly to environmental stimuli. Despite its small size and low cost, the uBot-4 has the potential to perform many dexterous tasks. Figure 8 shows the current version performing shoveling, stacking, pushing and throwing tasks. For residential health care applications, we predict that the vast majority of initial behavior will require acquiring objects, transporting, and then placing them in a new location—the so-called "pick-and-place" tasks. This is demonstrated in Figure 8 by bucket stacking, pushing, drawer pushing, and plowing.

Dynamically stable robots are well suited to human scale environments because they maintain postural stability, even given a small footprint. Active stabilization becomes easier as the robot (and thus the center of mass) becomes taller. Active postural control can also be used to generate greater pushing and pulling forces than are possible on an equivalent statically stable platform (Thibodeau, Deegan, & Grupen 2006).

4 Applications

The following are examples of innovative applications made possible by the introduction of a general purpose mobile manipulator designed for in-home environments, and the rest of the core technologies described in the previous section.

4.1 Family-Client Interfaces for Social Interaction: The Televisit

Isolation often leads to depression. Thus enhancing the ability to communicate is an effective way to improve the general well-being of an elder. The introduction of a mobile manipulator embedded with the capability to relay communication between the elder and the outside world brings much more flexibility to the application. The elder can now interact with an embodiment of the remote family member, rather than just listening to their voices over the phone. Authorized friends and

Figure 8: The uBot-4 performing plowing, stacking, pushing, and throwing tasks.

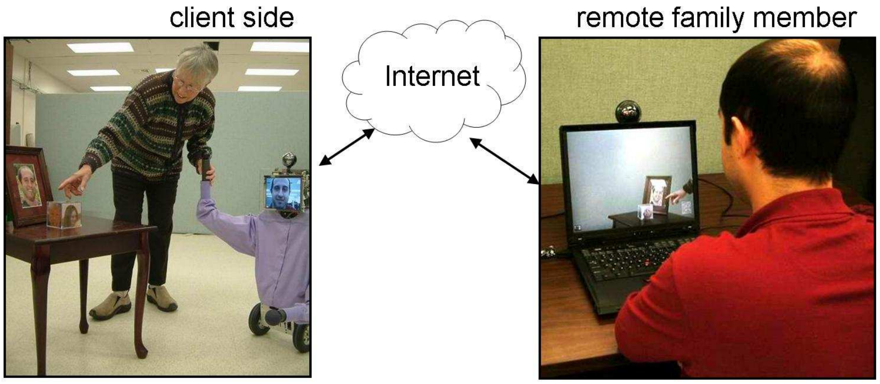


Figure 9: A client gives a tour of the assisted care facility. The visitor and the host share a videophone conversation while interacting and moving about the facility.

family can now move about the residence and interact with the client-side environment. This enables the family member to maintain a presence in the assisted care facility—elders and family members can engage in joint tasks that are more fruitful and enjoyable than stationary visual and auditory interfaces.

Figure 9 presents an example in which a client hosts a visit and gives a guided tour of the facility to a remote friend that teleoperates the uBot. The client can lead the robot, point at objects, etc. During such visits, the uBot may actually perform physical work such as picking up objects from the floor or doing other chores with the client or independently. We hope to create a social experience for the senior, and to make the virtual visit experience fun for the remote family member, like a gaming experience. In principle, some clients can use the uBot in this manner as well, to go places where their disabilities may not permit them to go otherwise. We intend to test to what extent we can meet these goals.

4.2 Community Services Interfaces: Emergency Service

Falls are the leading cause of injury-related visits to emergency departments in the United States and the primary cause of accidental deaths in persons over the age of 65 years (Fuller 2000). The statistics are frightening. More than one third of persons 65 years of age or older fall each year; in half of such cases the falls are recurrent (Tinetti 2003) and 60% occur at home. Furthermore, 95% of hip fractures are caused by falls, 40% of those hospitalized for hip fracture do not return to independent living, and 20% will die within a year (LifeLine 2006). Therefore, it is a major concern for elders who are at risk for losing the ability to maintain independent living.

Fall detection in this work is accomplished by modeling activities using Hidden Markov Models (HMMs) (Rabiner 1989) and the tracking features provided by the distributed camera array (Section 3.1). Training data for a "fall" activity is collected by having a person perform several different common actions such as walking and sitting, along with simulated falls in view of each camera. Position, velocity, and size information is provided

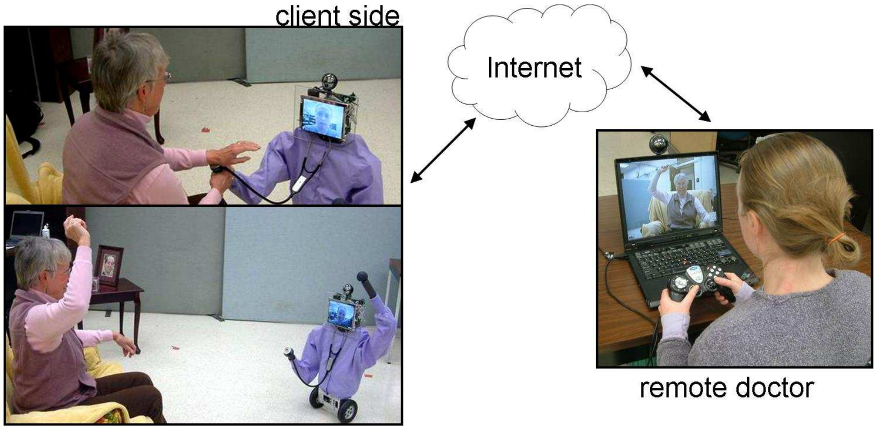


Figure 10: A three-question stroke diagnosis through video phone with motor tasks that are demonstrated by the uBot. Heart rate and respiration telemetry is also relayed to the doctor.

by the localization service at each video frame to create a sequence of feature points for each instance of an action. Two HMMs are trained using this data and the EM algorithm (Dempster, Laird, & Rubin 1977). One HMM is trained on the non-fall activities, and the other just on falls. When a new activity sequence is observed by the sensor array, it is evaluated under each of the two HMMs and classified as fall or no-fall based on the HMM that yields the maximum log likelihood for this sequence. The system constantly monitors a person's current activity for evidence of a fall, and can dispatch an alert if one is detected.

4.2.1 Fall Event: Responsive Client Receives a Virtual House Call

According to the National Institute of Neurological Disorders and Stroke, about 700,000 people have strokes each year. A simple diagnostic procedure can help to rule out stroke and if applied rapidly after a suspicious event, can make the difference between full recovery and debilitating and permanent brain trauma. A simple diagnostic test involves asking the subject to smile (exercising facial motor systems), to raise both arms (exercising areas of the brain associated with gross motor functions), and speaking a simple sentence (to engage linguistic centers).

If the subject fails to respond to any of these requests, prompt intervention can make all the differenece. Figure 10 presents the case when a client has fallen and is responsive. In this situation, the uBot and the sensor array initialize a virtual house call where the client, for example, undergoes such a diagnostic examination. The uBot is shown demonstrating an arm movement to the client under the direction of a remote doctor. The potential to act immediately even if the event occurs without other human observers is a potentially life-saving feature. Moreover, this intervention is a significantly more efficient use of the medical infrastructure than the alternative of making an appointment several days later and then examining a client who may be undiagnosable by then or for whom the damage is already done.

4.2.2 Fall Event: Unresponsive Client

Suppose that our hypothetical client has fallen and is unresponsive. We have prototyped an application in which the fall event triggers an automated attempt to rouse the client, places phone calls to emergency contacts (family member, emergency medical technician, or a 911 operator), and summons an ambulance. The uBot could respond either as a device that is teleoperated by an emergency contact, or autonomously if teleoperation is not available. Figure 11 illustrates the uBot approaching the client and applying a digital stethoscope for the purpose of providing the emergency medical technician enroute with telemetry that may aid the treatment of the subject.

Performing remote physical diagnosis, triage, or ruling out false alarms are all possible if the distributed sensor network includes a mobile manipulator such as this. A variety of instrumentation can be on hand depending on the needs of the client. Again improved response times, and minimal impact of medical service providers are the key performance criteria with which to judge the effectiveness and impact of this class of technologies.

4.3 Autonomous Assistance in Daily Living: Fall Prevention

In the previous scenarios, the uBot is controlled by a remote operator. As part of the distributed sensor-effector network, the uBot is also capable of acting autonomously as demonstrated in the following example.

A person's environment can contain risk factors that increase the chance that a fall will occur. A common

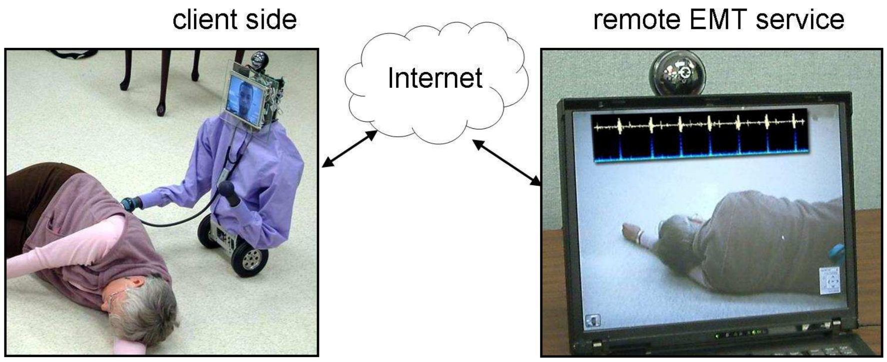


Figure 11: A fall is detected, the system enables access of an EMT. Using the uBot, the EMT can attempt to rouse the client, or in this case, apply a digital stethoscope, which relays heart rate and respiration telemetry to responders.

example is an object that unexpectedly ends up on the floor in a high traffic area, creating a tripping hazard. The risk that such objects pose can be eliminated by detecting and removing them before an accidental fall takes place. To our knowledge, no assistive technology has been demonstrated that can proactively decrease the probability of falling. However, in our framework, this application becomes possible since an embodied agent in the form of a mobile manipulator is considered as part of the distributed sensor-effector network.

Heavily trafficked areas in an environment can be determined using the HMM-based activity modelling approach described in Section 3.1.2. The floor plan view on the top left of Figure 12 shows high traffic areas in the hypothetical eldercare facility. The distributed array of camera sensors continuously monitors these regions for activity. Activities here are defined as *noticeable* temporal changes in the image obtained from the camera network. Such an event can however be triggered by human movement, changes in lighting conditions as well as the appearance of other possible objects in these trafficked areas. Hence, further image processing is done on the occurrence of such an event to determine if the *object* is in fact an obstacle. Two of these camera sensors then collaborate to localize this object. This global position acts as a goal to a Harmonic path planner (Connolly, Burns, & Weiss 1990) for a mobile manipulator. This can be seen in second row of Figure 12 where the dropping of a box by a delivery person in an active client region causes the sensor array to compute a path planner to move the obstacle away. The mobile manipulator is then guided autonomously by these camera sensors by providing motor commands to move the robot towards the object and noting the change. The cameras update the global position of the robot at every time step and feed this into the planner to obtain updated motor commands. This is shown in the third row of figures where the left image corresponds to the manipulator's estimated position in the planner and the right image corresponds to the actual position. When the manipulator determines that it is in contact with the object, it starts applying force to the object in order to push it away from the region of high activity, as is shown in the last row of Figure 12.

5 Conclusion

We have presented a general purpose mobile manipulator as part of a distributed sensor-effector network for applications in assistive living. In particular, the mobile manipulator adds new dimensions in our application design due to enhanced mobility, improved physical presence, and the ability to perform work. As a result, a wide range of cost-effective eldercare solutions can be devised, several of which are introduced in this paper. Qualitatively, these demonstrations show promising results in improving family-client interfaces for social interaction, interfaces to community services and medical providers, and autonomous assistance in daily living. The overall closed-loop developmental framework (ASSIST) that enables continual interaction between technological innovators, service industries, social workers, and the elderly clientele, is also presented. The quantative assessment of the marginal impact of these new technological innovations in the lives of real clients, is the next step in our development cycle.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant No. SES-0527648 and Grant No. IIS-0535120. The work is also supported under Grants ARO W911NF-05-1-0396, Army Research Office, Discrete Mathematics and

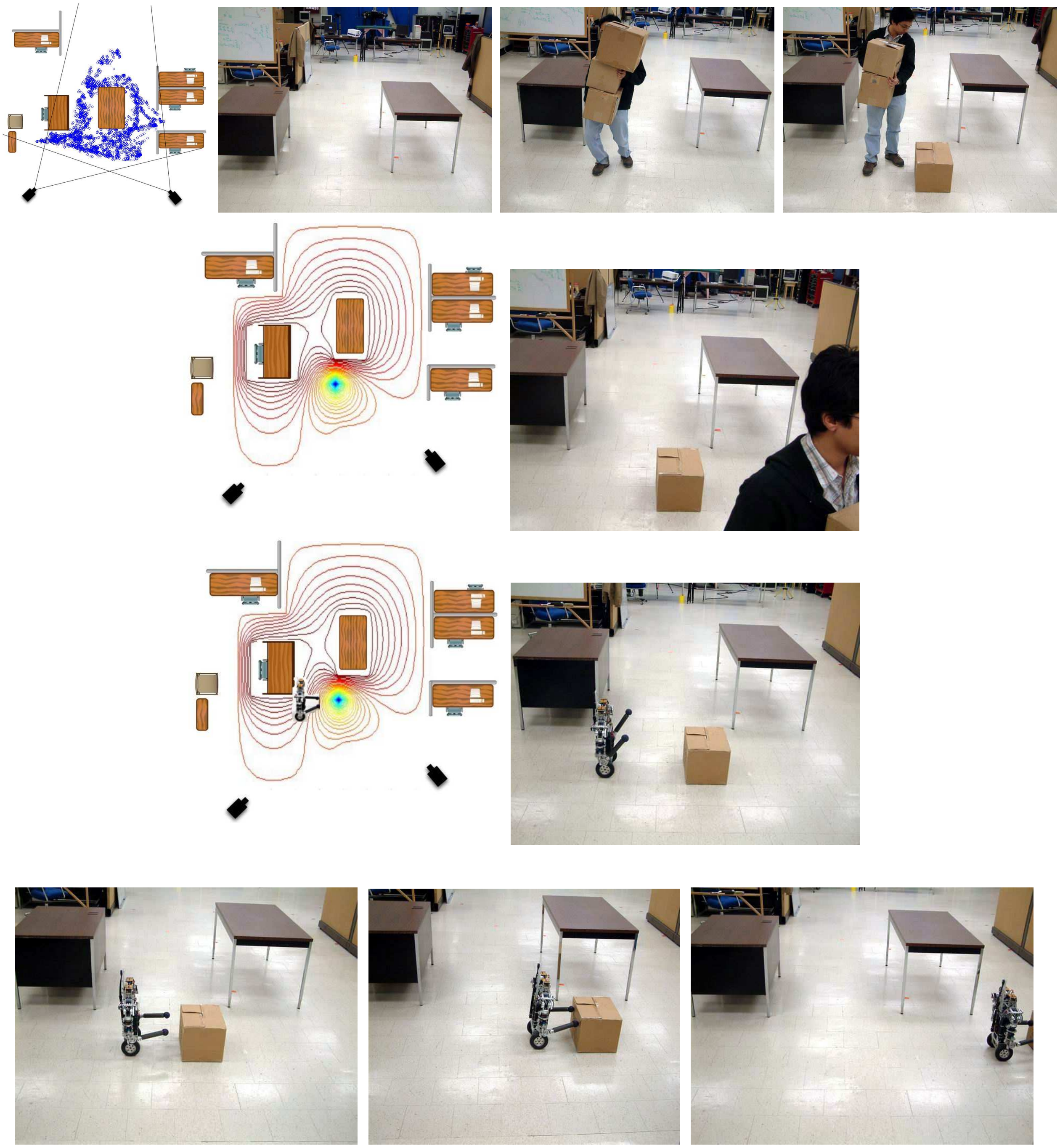


Figure 12: A delivery person drops a box in an active client corridor in a residential eldercare facility. The box is automatically detected since an unusal object appears in a high traffic pathway (the left-most image of the first row). A plan for removing the obstruction is constructed (left image in the second row), and the ASSIST mobile manipulator pushes the package out of the activity corridor.

Computer Science Program and NASA NNJ05HB61A, NASA Human and Robotics Technology Program, subcontract number 5710001842.

The authors wish to thank our social science colleagues at the Smith College School for Social Work for their contributions to the development of the ASSIST framework. In particular, the result of focus groups involving representation of the senior community, their caregivers, their family, and health care agencies has been of great importance.

References

Burgar, C. e. a. 2000. Development of robots for rehabilitation therapy: The Palo Alto VA/Stanford experience. *Journal of Rehabilitation Research and Development* 37(6):663–673.

Burrell, J.; Brooke, T.; and Beckwith, R. 2004. Vineyard Computing: Sensor Networks in Agricultural Production. *Pervasive Computing* 3(1):38–46.

Connolly, C. I.; Burns, J. B.; and Weiss, R. 1990. Path planning using Laplace's equation. In *Proceedings of IEEE Conference on Robotics and Automation*, volume 3, 2102–2106.

Czaja, S. J., and Sharit, J. 2003. Practically relevant research: Capturing real world tasks, environments and outcomes. *The Gerontologist* 43(1):9–18.

Czaja, S. J.; Sharit, J.; Charness, N.; Rogers, W. A.; and Fisk, A. D. 2002. CREATE: Center for Research and Education on Aging and Technology Enhancement. In Pieper, R.; Vaarama, M.; and Fozard, J. L., eds., *Gerontechnology: Technology and aging - Starting into the third millennium*. Aachen: Shaker Verlag. 54–61.

Dempster, A.; Laird, N.; and Rubin, D. 1977. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B* 39(1):1–38.

Eschenauer, L., and Gligor, V. D. 2002. A Key-Management Scheme for Distributed Sensor Networks. In *Proceedings of the 9th ACM conference on Computer and communications security*, 41–47.

Fuller, G. 2000. Falls in the elderly. *American Family Physician* 61(7):2159–68, 2173–4.

Ganesan, D.; Govindan, R.; Shenker, S.; and Estrin, D. 2001. Highly-Resilient, Energy-Efficient Multipath Routing in Wireless Sensor Networks. In *Mobile Computing and Communications Review*, volume 4(5).

Gist, Y. I., and Hetzel, L. 2004. We the people: Aging in the U.S. *U.S. Census Bureau 2000, special reports, U.S. Census Bureau series LENSR-19*.

Haigh, K. Z.; Kiff, L. M.; and Ho, G. 2006. The Independent LifeStyle Assistant: Lessons learned. *Assistive Technology* 18(1):87–106.

Heerink, M.; Krose, B. J. A.; Wielinga, B. J.; and Evers, V. 2006. Human-Robot User Studies in Eldercare: Lessons Learned. In *Proceedings of International Conference on Smart Homes and Health Telematics*.

Hobbs, F., and Stoops, N. 2002. Demographic trends in the 20th century. *U.S. Census Bureau, Census 2000, special reports, series CENSR-4*.

Karuppiah, D.; Grupen, R.; Hanson, A.; and Riseman, E. 2005. Smart resource reconfiguration by exploiting dynamics in perceptual tasks. In *Proceedings of International Conference on Robotics and Systems (IROS)*, 1513–1519.

Kassner, E. 2006. Home and Community-Based Long-Term Services and Supports for Older People. *AARP Public Policy Institute*.

Koester, H. H. 2004. Usage, performance, and satisfaction outcomes for experienced users of automatic speech recognition. *Journal of Rehabilitation Research and Development* 41(5):739–754.

Kotlikoff, L.; Fehr, H.; and Jokisch, S. 2005. Aging, the world economy and the coming generational storm. *National Center for Policy Analysis Study*.

Kumar, V.; Rus, D.; and Singh, S. 2004. Robot and sensor networks for first responders. *Pervasive Computing* 3(4):24–33.

Lachman, M., and Andreoletti, C. 2006. Strategy use mediates the relationship between control beliefs and memory performance for middle-aged and older adults. *The Journal of Gerontology Series B: Psychological Sciences and Social Sciences* 61:88–96.

Leifer, L. 1981. Rehabilitative robots. *Robotics Age*.

Li, D.; Wong, K.; Hu, Y. H.; and Sayeed, A. 2002. Detection, Classification and Tracking of Targets in Distributed Sensor Networks. *IEEE Signal Processing Magazine* 19(2):17–30.

LifeLine. 2006. The lifeline medical alert service.

Mainwaring, A.; Polastre, J.; Szewczyk, R.; Culler, D.; and Anderson, J. 2002. Wireless Sensor Networks for Habitat Monitoring. In *WSNA*, 88–97.

Martinez, K.; Hart, J. K.; and Ong, R. 2004. Environmental Sensor Networks. *Computer* 37(8):50–56.

Michael, Y. L.; Colditz, G. A.; Coakley, E.; and Kawachi, I. 1999. Health behaviors, social networks and healthy aging: Cross sectional evidence from the Nurses' Health Study. *Quality of Life Research* 8(8):711–722.

Morris, A.; Donamukkala, R.; Kapuria, A.; Steinfeld, A.; Matthews, J. T.; Dunar-Jacob, J.; and Thrun, S. 2003. A Robotic Walker that Provides Guidance. In *Proceedings of IEEE International Conference on Robotics and Automation*, volume 1, 25–30.

Noury, N.; Herve, T.; Rialle, V.; Virone, G.; Mercier, E.; Morey, G.; Moro, A.; and Porcheron, T. 2000. Monitoring behaviour in home using a smart fall sensor. In *IEEE-EMBS Special Topic Conference on Microtechnologies in Medicine and Biology*, 607–610.

Pin, S.; Guilley, E.; Spini, D.; and Lalive d'Epinay, C. 2005. The impact of social relationships on the maintenance of independence in advanced old age: findings of a Swiss longitudinal study. *Zeitschrift fur Gerontologie und Geriatrie* 38(3):203–209.

Pollack, M. E. 2002a. Planning Technology for Intelligent Cognitive Orthotics. In *6th International Conference on AI Planning and Scheduling*, 322–331.

Pollack, M. E. e. a. 2002b. Pearl: A mobile robotic assistant for the elderly.

Rabiner, L. R. 1989. A tutorial on Hidden Markov Models and selected applications in speech recognition. *Proceedings of the IEEE* 77(2):257–286.

Rogers, W. A., and Fisk, A. D. 2004. Psychological science and intelligent home technology. *The Psychological Science Agenda* 18:5–7.

Sakagami, Y.; Watanabe, R.; Aoyama, C.; Matsunaga, S.; Higaki, N.; and Fujimura, K. 2002. The intelligent ASIMO: System overview and intergration. In *International Conference on Intelligent Robots and Systems*, 2478–2483.

Stiehl, W. D. e. a. 2006. The huggable: A therapeutic robotic companion for relational, affective touch. In *Proceedings of IEEE Consumer Communcations and Networking Conference (CCNC2006)*.

Tanaka, F., and Suzuki, H. 2004. Dance interaction with QRIO: a case study for non-boring interaction by using an entrainment ensemble model. In *Robot and Human Interactive Communication, ROMAN*.

Tapus, A., and Mataric, M. 2006. Towards socially assistive robotics. *Journal of the Robotics Society of Japan (JRSJ)* 24(5).

Tapus, A.; Mataric, M.; and Scassellati, B. 2007. The grand challenges in socially assistive robotics. In *IEEE Robotics and Automation Magazine (RAM), Special Issue on Grand Challenges in Robotics*.

Tapus, A.; Tapus, C.; and Mataric, M. 2007. Hands-off therapist robot behavior adaptation to user personality for post-stroke rehabilitation therapy. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), Roma*.

Thibodeau, B. J.; Deegan, P.; and Grupen, R. 2006. Static analysis of contact forces with a mobile manipulator. In *Proceedings of the 2006 IEEE International Conference on Robotics and Automation*.

Tinetti, M. 2003. Preventing falls in the elderly. *New England Journal of Medicine* 348:42–49.

Williams, L.; Fels, D.; Treviranus, J.; Smith, G.; Spargo, D.; and Eagleson, R. 1998. Control of a remote communications system by children - a user-centered design methodology. *International Journal of Industrial Ergonomics* 22(4):275–283(9).

Williams, A.; Xie, D.; Ou, S.; Grupen, R.; Hanson, A.; and Riseman, E. 2006. Distributed smart cameras for aging in place. In *ACM SenSys Workshop on Distributed Smart Cameras*.

Zarit, S., and Zarit., J. 1998. *Mental disorders in older adults*. New York: Guilford Press.