Human-Robot Interaction in Service Robotics

H. I. Christensen*, H. Hüttenrauch†, and K. Severinson-Eklundh†
*Centre for Autonomous Systems †Interaction and Presentation Lab.
Numerical Analysis and Computer Science
Royal Institute of Technology
SE-100 44 Stockholm, Sweden

{hic,hehu,kse}@nada.kth.se

Abstract

Service robots or personal robots are to be used by non-expert users. This requires special attention to the human-robot interaction to provide systems that people are comfortable to use and be around. In this paper a number of different modalities for human-robot interaction are presented and a few systems that exploit such interfaces are outlined. These systems have been used for initial studies of use with non-expert users, the results are briefly summarised and issues for future research are outlined.

1 Introduction

The western world is at the verge of major changes in terms of elder care. Over the next 5-10 years the number of elderly people in need of care[1] will increase by 50%. At the same time the number of retired people will increase by more than 50% over the next 25 years[2]. To maintain an independent style of living and a level of autonomy, while limiting social costs, there is a need for service robots or personal robots that can assist elderly in their homes. Typical tasks for such systems include helping people getting dressed, simple cleaning of the house, providing assistance for mobility (leading a hand), preparing food, etc. Such robot systems will not only be of utility to elderly. On the short term it can be expected that such robots also will be of significant utility to handicapped, and as prices come down the market will gradually expand to include the regular household, and corporate care institutions.

A significant problem for the deployment of such robotics systems is design of user interfaces that enable non-expert users such as the average elderly citizen to interact and programme these systems. Design of the user interface is thus key to the acceptance of such systems. Traditional industrial robots are used by trained operators that have a strong model of the operation and limitations of the system. Both for traditional manipulators and for mobile platforms such as autonomous guided vehicles (AGVs) the programming is carried out using traditional programming languages or through graphical user interfaces that provide a well-structured dialogue. Such a user interaction model is not particularly suited for systems that are to be operated

[1]In terms of elder care it is widely recognized that most people older than 85 are in need of assistance to maintain a respectable quality of life

[2]These data are based on material from the Swedish Bureau of Statistics, but are valid for most of Europe, USA and Japan

by non-experts. For such systems there is a need for more flexible and more intuitive interaction, or collaboration between the user and the robot.

In this paper a number of different modalities for interaction with users are presented and an overall model for non-expert user interaction is discussed. A prototype system for a test application is presented and discussed in detail. Results from an initial user study are also presented. Finally the overall issues of human-robot interaction are summarised and issues for future research are presented.

2 Interfaces for service robots

In order for regular citizens, i.e. non-experts, to interact with robots there is a need for highly flexible interfaces. A number of obvious options can be considered. To provide a non-expert view on the utility of different modalities, a small survey of potential users (134 representative people) was conducted (1). In this study people were asked to select the modalities to be used for interaction, the choices were not exclusive, the following results were reported: 82% preferred speech, 63% a touch screen, 51% gestures, 45% some command language. Speech was thus by far the most popular choice.

In the following we will briefly consider pros-and-cons for a number of input and output modalities to give an impression of the choices available.

2.1 Entering commands

Some of the obvious choices for entering commands to a robot are:

1. Touch Screen
2. Spoken Dialogue
3. Gestures
4. Compliant Guidance
5. External Interfaces

Using the traditions from regular factory floor automation the touch screen is a natural choice. This setup was for example used in the widely cited RHINO robot from Bonn, that did tours of Museums. The advantage of a touch screen is that it is easy to present the possible alternatives and the input from the user is unambiguous. The by far major disadvantage of a touch screen is that the user has to be in front of it. For a mobile platform this is a significant constraint and in particular for feedback it is inconvenient to 'run' after the robot.

Speech is by far the most preferred input modality according to our user survey. Today it is possible to get fairly robust speech recognition systems for limited vocabularies (2) and for large vocabulary systems such as Dragon Dictate and IBM ViaVoice it is possible to obtain recognition rates of up to 95%. This is, however, for dictation, for interpretation of spoken commands it is significantly harder to achieve such levels of accuracy. It is thus possible to use speech for commanding systems as for example demonstrated by Sagerer et al (3). Unconstrained speech interpretation is however still very difficult. A problem with speech input is often the signal-noise ratio. E.g., in a room with several people the robot might not be able to distinguish the user from

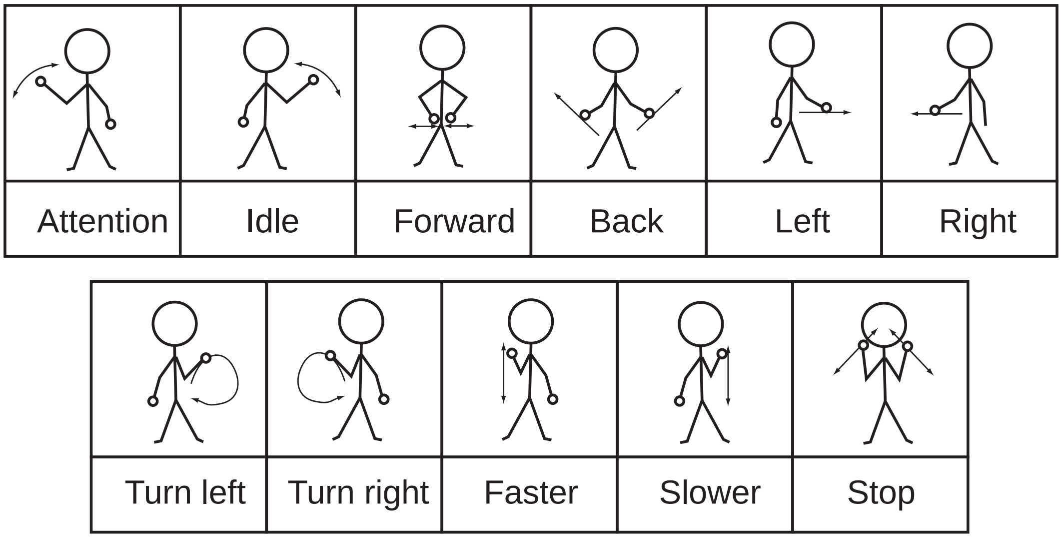

Figure 1: Example gestures for simple robot interaction

background noise and there might be a need to use a head-set to provide robust speech input.

Gestures is particularly useful in combination with speech to facilitate use of commands such as "pick *that* up", where "that" is specified by a pointing gesture. In addition waving gestures to get attention can be useful. Using colour vision in combination with tracking it is possible to track the hands and the head of a person, and use these for mixed dialogues. Such work has been reported by Cipolla et al (4), Swain et al (5). Both of these systems use simple pointing to specify the position of objects. More advanced user interaction has been explored by Sandberg (6). In this work Markov Models are used for recognition of gestures. Examples of possible gestures are shown in figure 1. So far gesture recognition is also a non-robust input modality. Variations in background, lighting, etc. makes it challenging to provide robust input.

To provide robust input to the user for speech and gestures it is thus critical to integrate these modalities with a dialogue system that allow recovery from partial interpretation of user commands.

All of the modalities mentioned so far are non-contact. For instruction of the robot to carry out specific actions it is not always convenient to use non-tactile interfaces. It is for example useful to be able to guide the robot to open doors, to guide it to places. I.e., for going about the house it is convenient to "take its hand" and lead it around the house for initial map acquisition. For this is is convenient to use force-torque sensing at the tip of a manipulator. This poses however a significant demand on the control of the robot as compliant guidance of the robot as is know from "teach-in" systems (7), poses significant safety demands on the robot to ensure compliant and safe operation in a natural environment. Building compliant manipulators for operation in environments inhabited by human is a tough problem as for example demonstrated by Hirzinger et al. (8).

The above examples have all assumed that the user is interacting directly with the robot, an alternative that is getting more and more popular is use of Personal Data Assistants, such as the Palm computer. Most PDA's have an IR link that can be used for communication with the service robot. Using such devices it is possible to display a map and point to places on the map to direct the robot to pre-defined locations. It is here even possible to combine a remote unit with a touch screen interface. Another examples might be use of a WAP enabled cellular phone, or a device similar to an advanced remote control unit for home entertainment.

2.2 Providing feedback

In terms of providing feedback to the user, several obvious aspects should be considered:

1. Graphics
2. Audio/Speech feedback
3. Self-evident motion
4. Animated characters

Graphics is a convenient way to illustrate maps and potential actions. Graphics provides both context and a fixed set of options. The disadvantage is that presentation of graphics might require the user to be close to the robot for receive the output. For portable devices such as PDA's small characters and maps might be difficult to read in particular for elderly people or visually impaired. The modality should thus be used with caution.

Spoken feedback is useful as it provides a simple interface that is easy to understand and it can in particular be used to explain actions as they are executed. One problem with many low-end speech synthesis systems is, however, poor quality. The synthesized speech is clearly artificial and it might be difficult to understand for the casual users. In addition synthesis of good quality speech might not be available in all the needed languages. The best systems are at present only available in major languages like English. An important problem to consider is the ability to turn of spoken feedback. The feedback might be useful during training and initial execution, but after a while the feedback becomes a nuisance, and it is thus essential to provide a scalable feedback system that can be adjusted from verbose to mute.

To complement use of graphics and speech it is possible it use animated characters to provide intuitive feedback. One example of such a method is shown below in figure 2.

The character has four degrees of freedom. The head can nod and shake and both arms can be controlled. Thus when a command is entered the character might nod its head to indicate that it has understand the command and shake its head if it cannot carry out the commands, i.e. a yes and no response. In addition the head may be raised over the head to indicate a "lack of understanding" gesture that indicates that the commands was not understood. The use of such a character is highly intuitive and most people will have no problem understanding the feedback. This is thus a very strong cue for feedback in particular for non-expert users.

Experience from initial user studies clearly indicate that the operation of a service robot should be self-evident. It is critical to provide fast and consistent feedback to the user. Even small delays in the feedback result in uncertainty and it might not be obvious if the robot has understood an issued command. This is in particular true as part of the dialogue with a user (9). During execution of a particular mission it is also important to use self-evident motion for the robot so that the actions to be carried out by the robot does not surprise the users, as such 'surprises' will result in uncertainty (10).

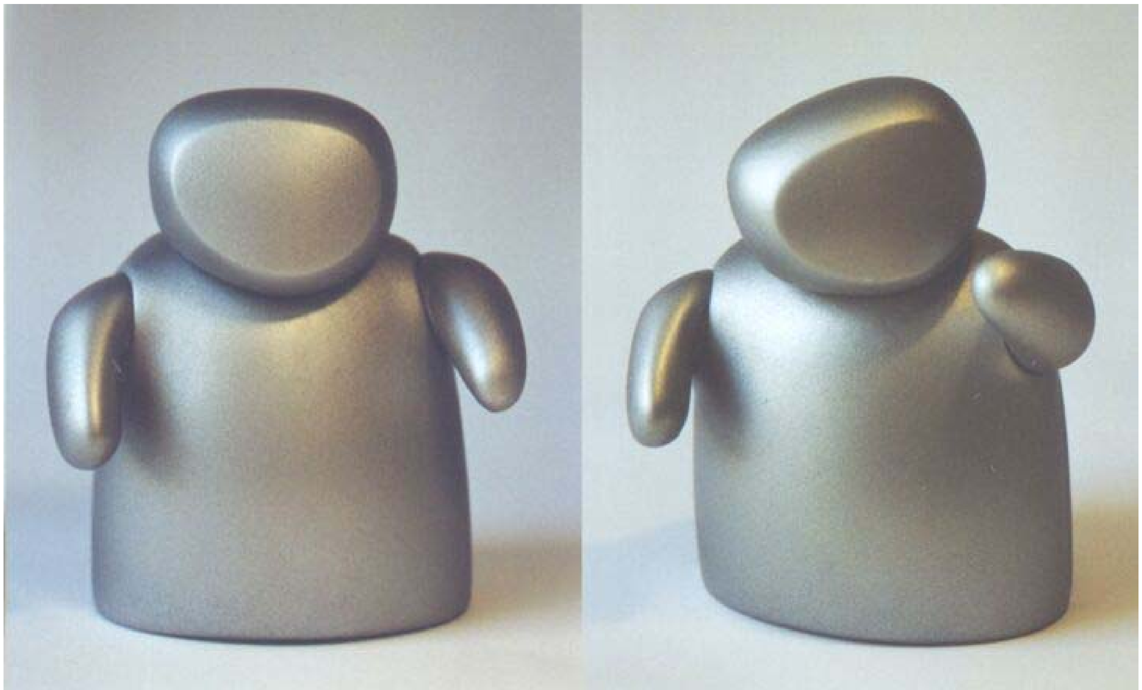

Figure 2: An example of an animated character that provides feedback to the user

3 Initial Systems

To investigate the use of the modalities outlined above a set of three different systems have been developed i) a delivery robot, ii) a prototype for a more competent service robot, and iii) an aid for handicapped. All there systems have been developed around Nomadic Technologies platforms. Example platforms are shown in figure 3

All these systems have been developed around a common architecture. The system is built around a hybrid-deliberative architecture, named BERRA (11, 12). The system is composed of a rich set of reactive behaviours for navigation, interaction, servoing on structures, etc. The behaviours are coordinated through deliberative control. The interaction with users and tasks is carried out by a joint interaction and planning module. The architecture is shown in figure 4. The system includes facilities for automatic map acquisition, and topological mapping. The test facility covers and in-door environment that is a 200 by 100 m office complex. Part of the environment is a laboratory that is setup as a realistic living room equipped with IKEA furniture.

The delivery agent and the service robot have speech, graphics, gesture, and PDA interfaces (IR link to a PalmPilot), while the handicap also includes the animated character mentioned earlier.

The systems have been in operation for about two years and carried out missions covering more than 100 km.

The robot can carry out tasks that involve:

- coverage of regions for cleaning
- fetch and delivery such as mail delivery in an office environment
- manipulation of objects, including opening and closing of doors
- person following (e.g. follow me)

Figure 3: A number of different platforms used for exploring service robot applications

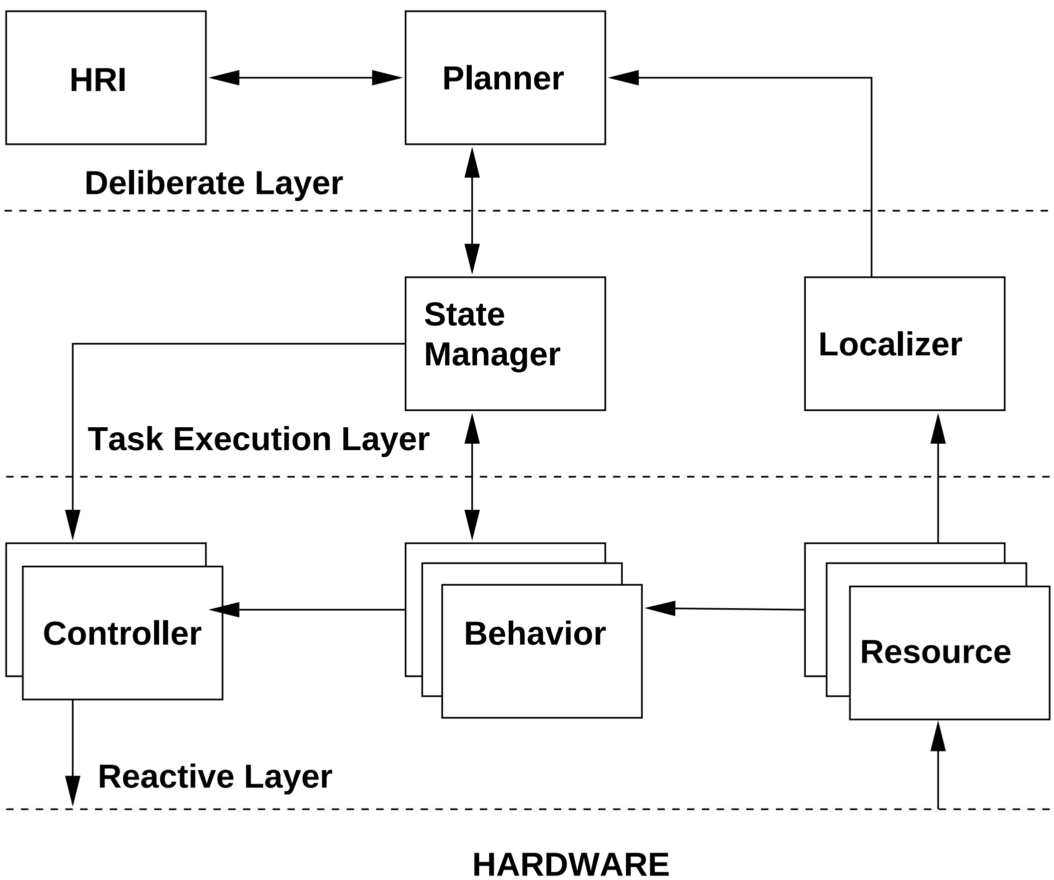

Figure 4: Architecture for the service robot systems

- guidance to visitors

4 User Studies

To investigate the use of these service robots, a set of initial user studies have been carried out (9). In these initial studies a person commanded the robot to go to a number of predefined locations. User studies that involve interaction have not yet been formalised.

The user study involved navigation and simple fetch-and-carry. As part of the study the test subject was asked to:

- Command robot to go to specific places
- Command robot to follow the subject

The subjects were then interviewed about their experience using the robot. The test subjects were all computer literature students that had no prior experience in robotics. The primary results from the study were:

- Even small delays cause uncertainty and the user has a tendency to have little patience with a robot.
- Design of behaviours must be such that motion is self-evident. The test-system used a highly conservative obstacle avoidance behaviour. Whenever an avoidance maneuver was initiated this surprised the user. Feedback is needed whenever a change of behaviour is initiated.
- Use of speech is natural for commanding the system and it is not until interaction is needed that the real value of gestures can be seen.

The user study involved primarily use of simple commands and we are thus not in a position to really evaluate the use of more interaction oriented modalities or feedback as part of mission execution.

The user study clearly demonstrated that an important aspect of a service robot is the overall design. In general the non-expert user has very high expectations with respect to the competence of a robot. The physical design strongly influences the perception of the robot and speech feedback results in a perception of "intelligence" that easily can be beyond the actual competence of the robot. It is thus essential to consider the overall physical design for a service robot as part of the process.

5 Conclusion and Future Work

In this paper a number of different modalities for human-robot interaction have been presented and a few experimental systems for evaluation of these modalities have been outlined. Some of these systems have been used in initial studies of use. State of the art is today such that deployment of advanced interfaces is possible. It is, however, essential that the design, implementation and deployment of systems is considered from a holistic point of view as all aspects from dialogue, modalities, physical design and actual control of the robot strongly influences the user reaction to the final systems.

So far service robotics has primarily been studied from a technical / robotics point of view. For the deployment of such systems is realistic environments it is, however, essential that human interaction and design is studied as well. There is thus a need for joint HCI - Robotics research to enable formulation of interaction systems that are suited for robot systems that are to be operated by non-experts.

Initial research has indicated the potential of such research and a few experimental systems has been deployed, but there is a need for more formal studies of usability in relation to use and coordination of different input and output modalities. The basic components are now available to support such research. In the immediate future we expect to deploy the mentioned systems is realistic user environments to provide the basis for such studies of human-robot interaction.

Acknowledgment

This research has been sponsored by the Swedish Foundation for Strategic Research. Partial support has also been received from the Swedish National Labour Board. This support is gratefully acknowledged. The BERRA software system was provided by A. Orebäck and M. Lindström. The animated character (CERO) was designed and manufactured by E. Espmark.

References

[1] Z. Kahn, "Attitudes towards intelligent service robots," Tech. Rep. TRITA-NA-P9821, Numerical Analysis and Computer Science, Royal Institute of Technology, Stockholm, Sweden, December 1998.

[2] G. A. Fink, C. Schillo, F. Kummert, and G. Sagerer, "Incremental speech recognition for multi-modal interfaces," in *IEEE 24th Conf. on Industrial Electronics*, (Aachen), pp. 2012–2017, September 1998.

[3] G. Fink, N. Jungclaus, F. Kummert, H. Ritter, and G. Sagerer, "A distributed system for integration speech and image understanding," in *Intl. Symp on Artificial Intelligence*, (Cancun, Mexico), pp. 117–126, 1996.

[4] R. Cipolla, P. Hadfield, and N. Hollinghurst, "Uncalibrated stereo vision with pointing for a man-machine interface," in *IAPR workshop on machine vision application*, (Tokyo), December 1994.

[5] R. Kahn, M. Swain, P. Prokopowicz, and R. Firby, "Gesture recognition using perseus architecture," in *CVPR96*, pp. 734–741, 1996.

[6] F. Sandberg, "Vision based gesture recognition for human-robot interaction," Master's thesis, Electrical Engineering and Information Technology, Royal Institute of Technology, Stockholm, Sweden, December 1999.

[7] S. Y. Nof, ed., *Handbook of Industrial Robotics*. New York, NY: John Wiley ans Sons, Inc, 2nd ed., 1999.

[8] G. Hirzinger, M. Fischer, B. Brunner, R. Koeppe, M. Otter, M. Gerbenstein, and I. Schäfer, "Advanced in robotics: The DLR experience," *International Journal of Robotics Research*, vol. 18, pp. 1064–1087, November 1999.

(9) L. Ostericher, H. Huttenrauch, and K. Severinsson-Eklundh, "Where are you going little robot? prospects of human-robot interaction," in *Computer – Human Interaction 1999*, (Pittsburgh, PA), August 1999.

(10) J. Reikki, J. Röning, S. Väyrynen, and V. Tornberg, "Gerobotics," in *First Europe-Jana Symposium on Human-Friendly Robotics* (G. Giralt and H. Inoue, eds.), (Tokyo), November 1999.

(11) M. Lindström, A. Orebäck, and H. Christensen, "Berra: A research architecture for service robots," in *Intl. Conf. on Robotics and Automation* (Khatib, ed.), (San Francisco), IEEE, May 2000.

(12) M. Andersson, A. Orebäck, M. Lindström, and H. Christensen, *Intelligent Sensor Based Robots*, vol. 1724 of *Lecture Notes in Artificial Intelligence*, ch. ISR: An Intelligent Service Robot, pp. 291–314. Heidelberg: Springer Verlag, October 1999.