IntellWheels MMI: A Flexible Interface for an Intelligent Wheelchair

Luis Paulo Reis[1,2], Rodrigo A. M. Braga[1,2], Márcio Sousa[1], and Antonio Paulo Moreira[1]

[1] FEUP – Faculty of Engineering of the University of Porto, Rua Dr. Roberto Frias, s/n 4200-465, Porto, Portugal
[2] LIACC - Artificial Intelligence and Computer Science Lab., University of Porto, Portugal
lpreis@fe.up.pt, rodrigo.braga@fe.up.pt, marcio.mc.sousa@gmail.com, amoreira@fe.up.pt

Abstract. With the rising concern about the needs of people with physical disabilities and with the aging of the population there is a major concern of creating electronic devices that may improve the life of the physically handicapped and elderly person. One of these new solutions passes through the adaptation of electric wheelchairs in order to give them environmental perception, more intelligent capabilities and more adequate Human – Machine Interaction. This paper focuses in the development of a user-friendly multimodal interface, which is integrated in the Intellwheels project. This simple multimodal human-robot interface developed allows the connection of several input modules, enabling the wheelchair control through flexible input sequences of distinct types of inputs (voice, facial expressions, head movements, keyboard and, joystick). The system created is capable of storing user defined associations, of input's sequences and corresponding output commands. The tests performed have proved the system efficiency and the capabilities of this multimodal interface.

Keywords: Multimodal Interface, Intelligent Wheelchair, Intelligent Robotics.

1 Introduction

Physical injuries occur frequently caused by accidents affecting the mobile capabilities of individuals, among other damages. Physical injuries could also be caused by medical conditions, like brain palsy, multiple sclerosis, diseases respiratory and circulatory diseases, genetic diseases or chemical and drugs exposition. Usually, the physical deficiency result on a limited control of some muscles of the arms, legs or face. It's very difficult to generalize physical deficiencies and each person has different symptoms and uses different strategies to deal with it. An example is the cerebral palsy, which concern with injuries on some brain areas responsible for the movement control, resulting on a difficulty that could be slight or cause total incapacity of moving the arms, legs or even talk. Two persons with brain palsy are different on each one's deficiency and degree of muscle control. Cerebral palsy as no cure but the effects could change with the age.

J. Baltes et al. (Eds.): RoboCup 2009, LNAI 5949, pp. 296–307, 2010.

Nowadays, society is more and more concerned with enabling handicapped persons to have an as independent life as possible. Wheelchairs are important locomotion devices for handicapped and senior people. With the increase in the number of senior citizens and the increment of people bearing physical deficiencies in the social activities, there is a growing demand for safer and more comfortable Wheelchairs and the new Intelligent Wheelchair (IW) concept was introduced. Like many other robotic systems, the main capabilities of an intelligent wheelchair should be: Autonomous navigation with safety, flexibility and capability of avoiding obstacles; intelligent interface with the user; communication with other devices (like automatic doors and other wheelchairs). However, most of the Intelligent Wheelchairs developed by distinct research laboratories, [6][10], have hardware and software architectures too specific for the wheelchair model used/project developed and are typically very difficult to configure in order for the user to start using them.

The Intellwheels prototype includes most of the typical IW capabilities, like facial expression recognition based command, voice command, sensor base command, advanced sensorial capabilities, the use of computer vision as an aid for navigation, obstacle avoidance, intelligent planning of high-level actions and communication with other devices. However the project is based on two main innovative ideas that will tackle the abovementioned IW problems. Firstly the Intellwheels project is based on a generic IW framework that enables easy development of new intelligent wheelchairs and IW control algorithms. The framework is flexible enough to enable easy transformation of commercial wheelchairs into intelligent wheelchairs with minor hardware changes and to enable the introduction of new modules and algorithms in the intelligent wheelchair. It includes a complete IW simulation module enabling to conduct virtual reality and mixed reality experiments.

The second innovation is concerned with the Intelligent Wheelchair command methodology that is based on a flexible multimodal interface. The wheelchair is commanded at a very high-level using a high-level command language based on simple commands such as “go to bedroom”, “wander”, “follow wall”. The commands are triggered by user selected input sequences using the multimodal interface. An input sequence may be something like “blink left eye” and then “say: go” or any given sequence of inputs coming from distinct input devices. The wheelchair enables the user to even use the same type of input sequences to select its preferred inputs/action association. In order to enable the user to start using the wheelchair, a simple patient classification module based on machine learning techniques is now under development. It will be capable of identifying the user basic capabilities and enable him to start using the wheelchair flexible multimodal interface in a straightforward manner.

This work focuses in the development of a user-friendly multimodal interface, which is integrated in the Intellwheels project. This paper presents the first prototype of the multimodal human-robot interface developed that allows the connection of several input modules, enabling the wheelchair control through flexible input sequences of distinct types of inputs (voice, facial expressions, head movements, keyboard and, joystick). The system created is capable of storing user defined associations, of input’s sequences and corresponding output commands. This interface can provide an interaction between the wheelchair environment and the input method, so that at any instance the input information can be analyzed and checked if it’s reliable, to assure the user safety.

The rest of this paper is organized as follows. Section 2 describes the concept of multimodal interface and indicates some of its desired characteristics. Section 3 describes the Intellwheels project and its main features and characterics. Section 4 describes the developed work reagrding the IW multimodal interface and section 5 describes the experiments performed and the results achieved. Section 6 presents the paper main conclusions and some pointers to future work.

2 Multimodal Interfaces

Generically an interface is an element that establishes a frontier between two entities. When an interface is used to assist in the Human-Computer Interaction it is called a user interface, being able to be graphical or command line based.

Most of the traditional graphical user interfaces are based in the WIMP (Window, Icon, Menu, and Pointing device) paradigm, which uses the mouse and keyboard as physical input devices to interact with the interface, for example to access information or accomplish any needed task.

An evolution to this paradigm and a way to create a more natural interaction with the user is the establishment of a multimodal interaction. This interaction contemplates a broader range of input devices such as video, voice, pen, etc, and so these interfaces are called Multimodal Interfaces.

A Multimodal Interface [1] “processes two or more user input modes – such as speech, pen, touch, manual gestures, gaze, and head and body movements – in a coordinated manner with multimedia system output. They are a new class of interfaces that aim to recognize naturally occurring forms of human language or behavior, and that incorporate one or more recognition-based technologies (e.g., speech, pen, vision)”. This type of interface can be used in several fields such as, for example, navigational devices – [2] and [3] – and health care solutions – [4] and [5].

Considering the purpose of this work the main aspects to consider should be the adaptability to users, usability and safety. These factors are determinative in a Multimodal Interface design, where subjective characteristics, like user satisfaction and cognitive learning, and user interaction depend on them. The adaptability to users is necessary so that the interface can be usable and understandable by any person, independently from his informatics knowledge and cognition. With the multimodal interaction between inputs comes a wider range of output control options and a complementarily between inputs.

The output control is achieved by the combination of several inputs, only being limited by the total number of inputs. As the interaction between the inputs can differ depending on the environment, this multimodality achieves a complementarily that when any input become less recognizable, it can be compensated by another, but this must be done being in mind the interface accessibility.

Finally, having in account the project enclosure, the multimodality must enable the access to any user, despite his deficiency. This shows the Multimodal Interface accessibility importance, so that if a user as any deficiency that suppress the use of one input, there is another that compensates this handicap [4].

Since this is a Multimodal Interface, it is necessary that this project allows a transparent and intuitive control of the Wheelchair and also a flawless input interaction.

This is achieved by the understanding of the user and inputs interaction. The inputs interaction is one of the key points of a multimodal interface, since it will be this interaction that will produce the desired output to the user. It is necessary the existence of a support for integrating any kind of inputs like: video, speech, handwriting, gestures, etc, but also this support must contemplate a robust processing of the inputs to fully recognize the user intentions.

The user interaction is another key point of a multimodal interface, if not the most important, so that the user can have and pleasant experience with the interface. It is necessary to consistently verify the disposition of every component of the interface so that the visual information and content can be easily accessed. Also it is needed to assure an intuitive interaction with the system, regarding the information about the available actions and how the user can interact with them.

Other factor is the interface output, which is divided in two parts: the processes concerned with the interpretation of the user inputs and processes regarding the correct visualization of the information given to the user about the system state.

3 Intellwheels Project

This Multimodal Interface is included in the Intellwheels Project, which main objective is to provide an intelligent wheelchair development framework to aid any person with special mobility needs.

This project encloses the prototype of an intelligent wheelchair, since its hardware structure to all software needed to provide a flawless control and experience to the user, being the hardware architecture shown in figure 1.

This architecture was created with the objective of being flexible and generic, so that it does not imply considerable modifications in the wheelchair structure [6].

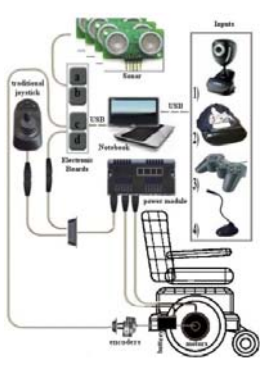

Fig. 1. Hardware architecture [6]

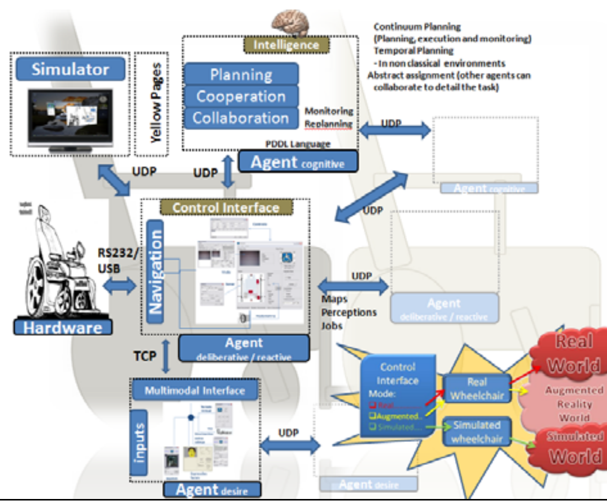

Fig. 2. Software architecture[6]

To enable a multimodal control of the wheelchair it is necessary to provide several inputs to user. It is also essential that these devices can map a broad kind of input methods, so that given any type of movement needs the user always has a way to control the wheelchair. With that in mind the following input devices were implemented: USB Joystick; Microphone; Head movements; Keyboard; Mouse; Video camera. With this it is possible to control the wheelchair using several types of inputs, from head movements to facial expressions [7] [8] [9]. Apart from the user inputs, the wheelchair also uses sensoring devices like: encoders, for the odometry calculation, and sonars, for obstacle detection. Several hardware interface modules are included to deal with the encoders and sensors.

One final, and important hardware device, is the laptop HP Pavillion tx1270EP, which is used to run all the developed software. In figure 2 is possible to see the global multi-agent software architecture defined for the Intelwheels project.

Focusing in the multimodal interface, it interacts with the Control Interface through a TCP socket connection, where the Control Interface will inform the Multimodal Interface of the available actions and state of any pending planning.

The user interacts with the Multimodal Interface which provides the connection, also through a TCP socket, of several independent input modules. The input modules are used for the user interaction and, therefore, create input sequences to execute the control actions assign by the Control Interface.

4 IntellWheels Multimodal Interface

The Multimodal Interface shows, in a graphical way, information about the actions, and input modules, such as kind, name or type of action or input, respectively. It also shows the defined input sequences, for the actions execution.

The joystick module works as a driver to establish a connection between an USB joystick and the Multimodal Interface. This module was adapted from [6], and it gets the information of the available buttons and analog sticks.

To enable the voice interaction it was necessary to implement a simple speech recognition module. The presented solution takes advantage of the IBM Via Voice [11] capabilities using the navigation macros, which allows the user interaction with any software through, previously recorded voice commands. However, the use of Via Voice has a disadvantage since it needs the voice module window to be active so that the voice commands macros can be perceived.

To assure the integration of the already developed inputs, the head movement module was adapted to communicate with the multimodal interface. This module takes advantage from one accelerometer installed in a cap, where it reads its values and transforms in a position type value, for pointer control, or in a percentage speed value to control the wheelchair.

4.1 Multimodal Interface Architecture

Since the wheelchair control platform and the multimodal interface are distinct agents, it was necessary to enable the multimodal interface agent to interact with the already developed control agent [6]. With that intention, data structure and information processing methods were created for the components interaction.

The system architecture, illustrated in figure 3, is a zoom in on the main architecture shown in figure 2. In this figure is possible to see the exchanged information between all the involved agents. The control interface acts as a communications server to the multimodal interface, as well as the multimodal interface acts as a communications server to the input modules.

Since the communications are totally established by the used Delphi components, as soon as the multimodal interface connects to the control agent, the control sends the information about the available high-level actions. For the input modules, as soon as one of them connects to the interface, firstly it sends its id and, number of module commands. Secondly, upon the receiving of a request from the interface, the input module sends the description of all the commands.

The interface information processing is divided in two logical parts: the server side and the client side processing. This division is derived from the need of the Multimodal Interface to act as a client to the control connection, but as a server to the inputs' connection.

For these models two data structures were created. One for storing the control actions information and other for storing the input modules commands information.

The inputs' structure is composed of six fields:

- Number: the internal number of the command;
- Name: the name of the command;
- Kind: this defines the name of the input module;
- State: for a button, this represents if its pressed – "True" – or if it was released – "False";
- Value 1 and value 2: these fields are used for transmitting the analog values of a command, for example the analog stick of the joystick.

For a digital command, like for example a button, the value fields will return a "n/a" string, being the same analogously applicable to an analog command, it returns the state field with a "n/a" string.

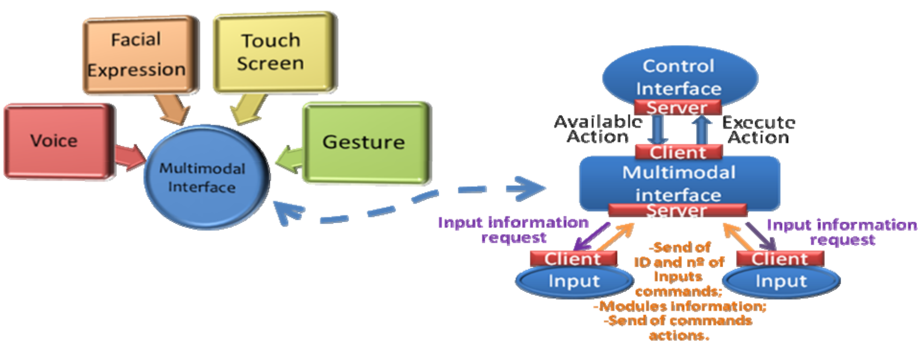

Fig. 3. Multimodal Interface architecture

The actions' structure is also composed of six fields:

- Name: the name of the action;
- Kind: the kind of action, for example movement;
- State: the availability for executing a action, returns "True" for a available action, or "False" if not available;
- Value: informs the interface about the execution of an action, returns "ON" if the action is under execution, or "OFF" when it stops its execution;
- Data: this field acts as an information about the level type of the action, being its options in table 1.

Table 1. Action structure: Data field

Data	Type name	Sent Parameters
0	Stop action	0
1	Manual action	2
2	Mid-Level action	1
3	High-Level action	0

The information is passed through one of the following messages:

- From the control interface:
```
<cmd id= name="" kind="" state="" value="" data=""\>
<cmd_state id="" state=""\>
```
- To the control interface
  - ➢ High level or Stop: `<action id=""\>`;
  - ➢ Mid level: `<action id="" value=""\>`;
  - ➢ Manual mode: `<manual value1="" value2=""\>`;
- From the input modules
  - ➢ Registration at the multimodal interface
```
<input_info id="" mods=""\>
<input id="">
<module num="" name="" kind=""\>
…
<module num="" name="" kind=""\>
<input\>
```
  - ➢ Input event generated by user interaction
```
<input_action num="" state=""\>
<input_action num="" value1="" value2=""\>
```

4.2 Input Sequences

The input sequences represent how the user interacts with the interface or being more precise how the user controls the wheelchair. These sequences are created through the combination of two or more input actions.

Independently from the input module kind, or if the command is digital or analog, the associated event has two common identifiers: the module id, and the command number.

To standardize the inputs events representation was defined that if the command is digital, then its state is "True" – T – when the button is pressed, or "False" – F - when the button is released.

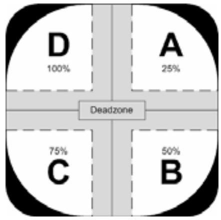

Fig. 4. Division for fixed speed values

For an analog command the state field is not used, but the input module will return the values of the analog axis. These are converted to a percentual value and are used to directly control the wheelchair, if the manual mode was activated, or are "processed" and define a fixed speed value.

This fixed speed value is achieved by logically "dividing" the cursor area in one of the analog axis of the joystick. Due to the short length of the stick it is only possible to divide in four areas without losing precision, being the division shown in figure 4. The variation assumes increments of 25% per zone, from A to D.

To simplify the sequence creation method, it was imposed a maximum number of six input command actions (fragments) to generate a sequence. Also, a minimum number of one input was imposed. Each fragment has the following format:

```
#<input_module_id>.<command_number><state>
```

The state field can be composed by one of three possibilities: "T" or "F", in case of a button, or "%" in case of one of the four pre established values (A, B, C or D).

The sequence entrance is limited by the already referred maximum number of fragments, or at any instance by the detection of an existent or nonexistent sequence. That is, for each fragment received the developed algorithm updates the input sequence under construction and, searches in the sequence list for the same occurrence. The search returns one of three available options:

- The occurrence is unique, and therefore the composed sequence can be immediately analysed;
- There are more occurrences of the same sequence fragment and thus it must be further processed;
- The occurrence does not exist in the list, meaning that the user is entering a not valid sequence and therefore the process is stopped;

If the search indicates the current sequence is not unique, the algorithm waits for a given predefined time for more input event actions in order to complete this sequence to a unique sequence. With this it is possible to evaluate if the user is still entering the sequence or, if during a pre established time interval none input action is received, if the sequence is already completed.

The use of this process turned the sequence input method more reactive to the user by providing an almost instant response to unique or wrong sequences, allowing a more effective control.

4.3 Interface Components

The interface components are all the interface visible components, since the simple buttons to images, menus and textual information. In order for the Multimodal Interface to be very simple it only contains the following components:

- List of available actions;
- Summary of the inputs connected;
- Input modules and control connections status;
- Input sequence graphical information;
- Sequence's list;
- Sequence's analysis result;
- Wheels speed information;
- Menus for programming the interface options and adding more sequences.

Fig. 5. Intelligent wheelchair multimodal interface design

All these components show the available information in a textual way, except the input sequence and wheels speed that show the information in a graphical way. Figure 5 displays the multimodal interface design.

5 Results

In order to evaluate the Multimodal Interface integration in the Intellwheels project several experiments were made using the Intellwheels simulator [12], with the objective of controlling the wheelchair in manual mode using distinct inputs.

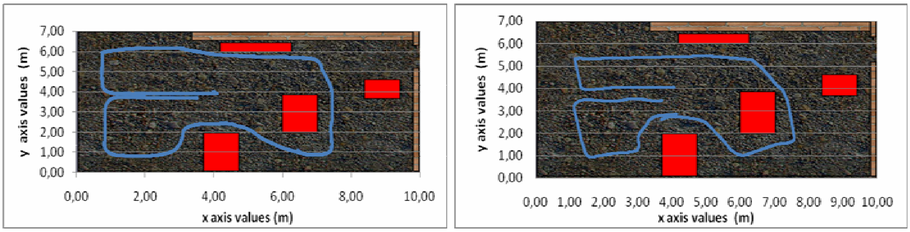

Fig. 6. Wheelchair movement in a room with obstacles, with joystick control and with head movement control

These experiments were made in a simulated room with several obstacles, where the wheelchair starts from a middle position and tries to go around the room perimeter, deviating from the objects, finishing in the start position. When using voice commands it was also tested the voice recognition software by introducing background noise during the tests. For the control with voice commands five commands (front, back, left, right and, stop) were defined to control the wheelchair. With these results of figure 6-a it is possible to see the wheelchair movement through the room, being the input method able to drive the wheelchair in the predetermined course without any problem.

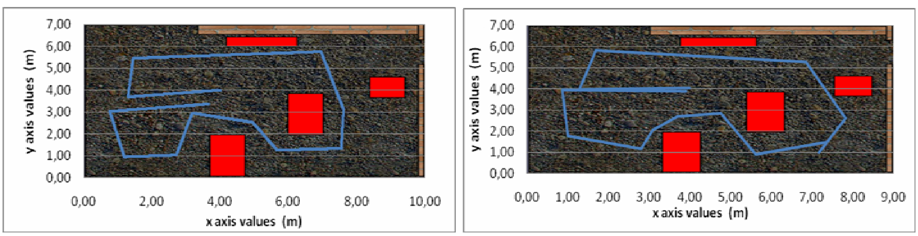

Fig. 7. Wheelchair movement in a room with obstacles, with voice control, without background noise and with background noise

Analysing the experiments of figures 7-a and 7-b it is possible to see that the implemented voice input method, for directly control the wheelchair, is a preferably input for open areas without obstacles. This is due to the delay in the response of the speech recognition software, which in emergency situations can become dangerous.

Another aspect to be considered analyzing these experiments is the sensibility of the speech recognition software to the background noise. During the experiments the microphone was approximately at 30 cm from the user, and the background noise source was a radio playing music with low volume. In these conditions it was necessary to repeat the voice commands several times, which has increased the experiment time and also made the control more difficult. Finally, figure 9 shows tests with a real wheelchair.

The results show that it is possible to drive the wheelchair just using head movements with good performance. However, these tests were made without any source of distraction. In other experiments it was confirmed that controlling the wheelchair

Fig. 8. Real wheelchair movement in a corridor with obstacles using joystick and voice control and using head and voice control

using only the head movement module, with a high sensibility, and with several sources of distraction is a very complex task.

Again, this movement method is preferred for open environments without obstacles. However the method is completely capable of manoeuvring the wheelchair in a crowded room, performing precise movement tasks, as long as the user has enough experience with this method.

Although the set of experiments performed was still very simple and separate simple experiments were performed for each input module, it is possible to take some interesting conclusions from the results. The inputs perform well and individually enable to control the wheelchair. However with distraction sources it is very complex to control the wheelchair with a single input and thus the use of high-level commands and input sequences to trigger them, seems to be an appropriate approach.

6 Conclusions and Future Work

The developed multimodal interface showed to be very flexible enabling the user to define distinct types of command sequences and associate them to the available high-level outputs.

To verify the system efficiency and the wheelchair control through the developed multimodal interface several experiments were conducted, where the wheelchair was controlled with the available inputs (joystick, voice, head movements and several inputs) in different kinds of environments (noise in the background, obstacles, etc.). The results achieved enabled to confirm the multimodal interface capabilities, except for the voice module, which proved not to be precise when there is noise in the background. However, the main capabilities of high-level commanding through input sequences of the multimodal interface need further experiments to be evaluated.

Some future directions for this project development are obvious and concern performing a set of experiments with the complete multimodal interface and the development of the yet missing input modules. One missing feature is a robust facial expressions recognizing module, needed to create a more multimodal experience to the user.

With the intention of making the Multimodal Interface more user friendly, a text to speech output and some kind of virtual user assistant could be implemented. These elements would function as an user integration process with the interface.

Acknowledgements

This work was partially supported by FCT Project PTDC/EIA/70695/2006 "ACORD: Adaptative Coordination of Robotic Teams" and LIACC/University of Porto. The second author would like to thank for CAPES for his PhD Scholarship.

References

1. Oviatt, S.: Multimodal Interfaces. In: Handbook of Human-Computer Interaction, pp. 286–304 (2002)
2. Johnston, M., Bangalore, S., Vasireddy, G.: MATCH: Multimodal Access to City Help. In: IEEE Workshop on Automatic Speech Recognition and Understanding, pp. 256–259 (2001)
3. Cohen, P.R., et al.: QuickSet: Multimodal Interaction for Distributed Applications. In: Proc. 5th ACM Int. Conf. on Multimedia, Seattle, Washington, United States, pp. 31–40 (1997)
4. Keskin, C., Balci, K., Aran, O., Sankur, B., Akarun, L.: A Multimodal 3D Healthcare Communication System. In: 3DTV Conference, May 7-9, pp. 1–4 (2007)
5. Zudilova, E.V., Sloot, P.M.A., Belleman, R.G.: A Multi-modal Interface for an Interactive Simulated Vascular Reconstruction System. In: Proceedings of Fourth IEEE International Conference on Multimodal Interfaces, pp. 313–318 (2002)
6. Braga, R.A.M., Petry, M., Moreira, A.P., Reis, L.P.: Intellwhels - A Development Platform for Intelligent Wheelchairs for Disabled People. In: 5th Int. Conf. on Informatics in Control, Automation and Robotics, Funchal, Madeira, Portugal, vol. I, pp. 115–121 (2008)
7. Martins, B., Valgôde, E., Faria, P., Reis, L.P.: Multimedia Interface with an Intelligent Wheelchair. In: Proc. of CompImage 2006 – Comp. Modelling of Objects Represented in Images: Fundamentals Methods and Applications, Coimbra, Portugal, pp. 267–274 (2006)
8. Faria, P.M., Braga, R.A.M., Valgôde, E., Reis, L.P.: Platform to Drive an Intelligent Wheelchair Using Facial Expressions. In: Proceedings 9th Int.Conf. Enterprise Information Systems - Human-Computer Interaction, Funchal, Madeira, pp. 164–169 (2007)
9. Faria, P.M., Braga, R.A.M., Valgôde, E., Reis, L.P.: Interface framework to drive an intelligent wheelchair using facial expressions. In: IEEE International Symposium on Industrial Electronics (ISIE 2007), pp. 1791–1796 (2007)
10. Simpson, R.C.: Smart Wheelchairs: A Literature Review. Journal of Rehabilitation Research & Development 42, 423–436 (2005)
11. IBM. Via Voice (June 2008), http://www-306.ibm.com/software/pervasive/embedded_viavoice/ (Consulted on July 2008)
12. Braga, R.A.M., Malheiro, P., Reis, L.P.: Development of a Realistic Simulator for Robotic Intelligent Wheelchairs in a Hospital Environment. In: Baltes, J., Lagoudakis, M.G., Naruse, T., Shiry, S. (eds.) RoboCup 2009. LNCS (LNAI), vol. 5949, pp. 23–34. Springer, Heidelberg (2010)