EasyHouse: A Multimodal Domotic System for the Tetraplegic

David M. Graça Tiago Guerreiro Joaquim A. Jorge
Department of Information Systems and Computer Science
INESC-ID/IST/Technical University of Lisbon
R. Alves Redol, 9, 1000-029 Lisboa, Portugal
david.graca@ist.utl.pt, tjvg@immi.inesc.pt, jorgej@acm.org

Abstract

Nowadays, we can find ourselves surrounded by technology, whether in public spaces, our homes or even within our body space. It is difficult to imagine how we could survive without some of the devices, and their functions, that we take for granted today but were not available a few years ago. However, some persons are deprived from this control which limits not only their movement and life quality but also the overall control over the surrounding environment. This document describes EasyHouse, a domotic multimodal platform for tetraplegics based on a matrix metaphor. The system uses gestures, classifying them into 8 basic movements, and speech as the main input modes. Results showed high user satisfaction and easy interaction while providing a low error rate.

Keywords

Assistive Technologies, Motor Disabled, Multimodality, Domotic System

1 Introduction

The World Health Organization (WHO) defines a person with disability (or activity limitation) as ,a person of any age who is unable to perform independently and without aid, basic human activities or tasks, due to a health condition or physical/ cognitive/ psychological impairment of a permanent or temporary nature [CEN 03]. According to the World Bank estimate about 10% of the world's population is suffering from some kind of disability. With these statistics we can change preconceived ideas about the disabled people because it proves that they're not an occasional exception, in fact they are a big piece of the global population that has serious needs which are frequently neglected and for whom little improvements can have significant results. This percentage will rise since globally the average age is rising, and it's believed that 58.5% of people over 65 have disabilities, with motor disabilities making up a large part of this.

In a house there are many appliances, most of them with different strategies for controlling the device, and almost none suited for the physically challenged. This situation imposes a dilemma for people that have a house and live in it, which is the lack of autonomous control over almost every appliance, transforming their house into a caregiver-operated house and not their own house. The focus of this work is on that subset of the tetraplegic users due to the serious needs that they have and the lack of existing usable solutions.

2. Related Work

This project uses a wide range of technologies, from domotic protocols to multimodal interfaces, therefore multiple areas of expertise had to be surveyed, to address the lack of mobility and user control of the motor-disabled. Many approaches have been developed, from Intelligent Robots to Smart Houses.

One important project was the ASPICE project that aimed at "*the development of the system that aid the neuromotor disabled to improve or recover their mobility (directly or by emulation)*" [Aloise 06] and had the goal of achieving maximum independence from the caregivers for its users. The main interface used was a Brain-Computer Interface (BCI) due to the lack of motor-skills requirements. The system used a graphical user interface as the user control unit. The GUI could control an AIBO robot [Cherubini 07] which could send messages to the caretaker such as "give me water" using it as a messenger. The main home appliances were integrated using a set of specialized actuators. For the interactions, all over the house, this project used 3 technologies: X10; IRRC; Wifi; depending on the type of device.

The Intelligent Sweet Home project is an ongoing project developed by the Korea Advanced Institute of Science and Technology (KAIST) for assisting the elderly and the handicapped based on 3 modules: Bed Robot System, Soft Remote Control and Network System. The **Bed Robot System**, can be separated in several sub-modules being the **Soft Remote Control System** [Do 05] responsible for the control of the system. It was responsible for recognizing the user intentions and transforming them into application commands. This Control System recognizes hand-motions and hand-postures using CCD cameras. Voice control was another possibility of this system that used simple commands, with a maximum of 2 words. The users controlled the surrounding environment by pointing to the desired device and issuing voice or gesture commands for confirmation. In the case of gestures the recognition rate was 95% in the recognition of a pointed object ranging the overall rate of 80,7% to 95,7% which was similar to a regular remote control using infrared waves. The success rate depended directly on illumination and normal moves like resting the hand could be interpreted by the system as commands. For voice recognition there were also some issues in noisy environments.

3. EasyHouse

We aimed at the development of an easy-to-use simple domotic interface that enhanced the users independence and relieved the caretakers workload. The system built covered many technological areas which had to be enclosed in particular modules. The system can be divided into 4 parts: Inputs, Outputs, Domotic Network and Core Operations:

Inputs is the module that deals with all the aspects of the interaction between the user and the system. In it there is a cursor tracking system that collects data and, when a movement is finished, transfers it to the recognition submodule. This submodule implements all the movement recognition algorithms tested. The input devices that use this module are the microphone, the laser beam and the touchpad.

Outputs implements both graphical, Fig. 1, and auditory feedback. Although two distinct modalities are used for feedback, they are triggered in different stages. The auditory feedback is triggered when an action occurs, emulating speech to describe the action triggered. The graphical feedback gives a visual overview of the actions detected and the current state of the house appliances in real-time. It provides feedback from both gestures and speech commands, reporting to the user both possible or accepted actions at that time.

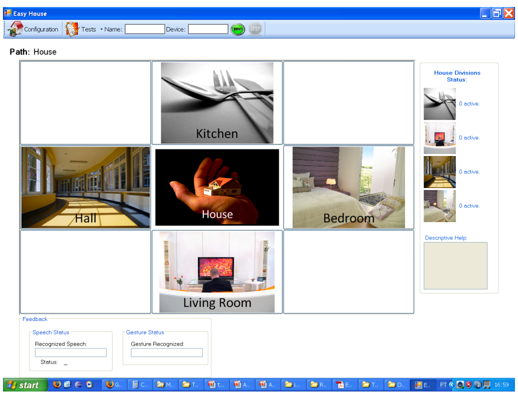


Figure 1. Graphical Interface.

Domotic Network is the module responsible for the actions in the household appliances. It receives commands from the system and reflects them in real actions in the household. It enclosures all the physical layer of the system, being the module of the system that delivers all the functionality to the users. In a house most devices remain in the same place for a period of at least 6 months but other devices can move between house divisions, so an easily adaptable domotic network was needed to fulfill the users needs. The X10 technology was chosen due to its wide availability, low cost and easy adaptation to a shifting environment. A C# .NET implementation of the X10 protocol was developed to create a Domotic network interface since available interfaces were not satisfying.

Core Operations are responsible for managing the data from the input devices and the feedback returned to the user. It represents the centerpiece of the system, controlling the graphical interface and analyzing the data sent from every input device available.

3.1. Gestures - Slash Interaction

The presented gesture approach is based on a Slash interaction, meaning that it uses low accuracy, fast movements as input. This approach was developed taking into account the users physical limitations, which are characterized by the lack of precision in small movements but relative constant direction in wider movements. To select options int he interface the movements are analyzed and a direction is found. For this study only 8 movements are considered due to their simplicity and reduced user motor-skills (vertical, horizontal and diagonal movements). The devices used control the cursor position, directly or indirectly, representing the cursor variations the data sent to analysis. This choice was taken due to modularity and compatibility issues since users capabilities are inconstant.

3.1.1. Gesture Recognition

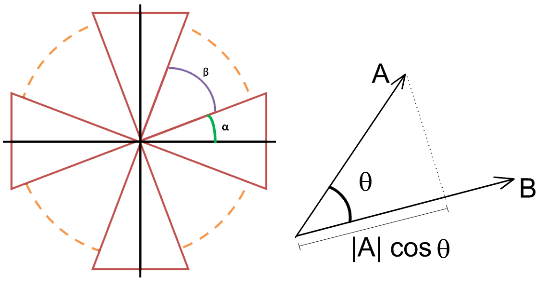
A gesture recognition algorithm, named **Discrete Algorithm**, was created based on a set of principles which took advantage of the lack of accuracy in initial movements. The users compensate the first stroke which might be inaccurate, needing to compensate that in a set of submovements that are directed to the correct position. An approach that simply uses that first and last position of the movement may detect with more precision the correct direction while having simpler calculations. The algorithm will then use those two points to create the direction vector which will then be classified according to the available directional classes. A angle is calculated between two vectors using F.J. Hill [Hill 00] formula, Figure 2(b).

Beyond these mathematical steps, the algorithm used two angles, the angles expressed the inclination limits, containing areas used for movement classification. α expressed the angle used for the vertical and horizontal movements and β expressed the angle for the diagonal movements, Figure 2(a).

Using suitable α and β angles, which were both 30º, and the slope formula, we were able to detect the direction of the movements in a precise and fast way.

3.2. Speech

Speech was a good candidate modality for the target population because all users had full control over speech and



(a) Gesture recognition algorithm based on the movement angles.

(b) Finding the angle between two vectors using a scalar projection.

Figure 2. Gesture Recognition Algorithm.

spoke English fluently. The high dimensionality and command rate together with the availability and level of user control made this modality the most compatible between users since. Moreover, as the system aims at in-house environmental control, the misrecognition due to a noisy environment is a smaller or not even an issue. There were 3 possible command types that addressed different contexts during the speech, each aimed at coping with different situations. The existing command types are Action, Navigation and Composite. As the names indicate, Action deals with device action, Navigation deals with the navigation commands going trough the house hierarchy and the Composite commands are a mix between the last two command types.

To rise the speech recognition rate and tackle the noisy environments problem, the dimensionality of the speech was reduced being constructed a valid grammar in a user-centered approach.

4. Evaluation

To assess the validity of the system, we have performed a set of evaluations with representative users from the target population. These evaluations are described in this section. The tests were performed in one session and each test was done individually. The task order and modality order varied from user to user, preventing biased results. In each modality or bi-modality test the first tasks presented were the even numbered tasks and when those are completed then the odd tasks are presented. The modalities/inputs used were, by order: Gestures using Touchpad, Gestures using Laser, Speech, Gestures using Touchpad + Speech, Gestures using Laser + Speech.

All users have regular upper body strength and coordination with the exception of the hands and arms, which they have a small degree of control and accuracy. These users have basic computer knowledge and are familiar with Microsoft Windows interfaces, using several devices or their own hands to make selections in the interface, Table 1.

There are several types of measures, captured automatically or manually, all without the user knowledge to prevent affecting performance. The automatically measures are Task Completion Times and Invalid Selections, while the manual measures are the Number of Helps and Users/Program Errors.

Table 1. Users technological profile table.

User ID	Age	Injury	Injury Type	Cause	Qualifications	Control Mechanisms
PF	31	C5	Complete	Dive	12th grade	Arm Stick
NC	34	C5	Complete	Fall	University Degree	Keyboard Buttons
HD	41	C6	Complete	Accident	12th grade	Hand/Arm Motions
SA	26	C4	Complete	Accident	University Degree	Eye Tracker

4.1. Accuracy Evaluation

There are two distinct analysis taken from this study, an analysis of the modalities and inputs, Interaction Analysis, and an analysis on the relation between number of commands of each task and the errors of the users, Command/Task Analysis.

Interaction Analysis. The interaction through the touchpad showed an average error rate of about 0,67% which was clearly inferior to all other devices. On the other hand, the laser was the device that performed the worst with a 9,52% error rate. As expected the speech showed some problems with a 6,47% error rate. However by mixing speech with the touchpad we improved the error rate by 13,79% when comparing with the only-speech approach. When mixing speech with the laser pointer we were able to reduce the error rate by 18,75% when comparing with only the laser pointer, see Figure 3.

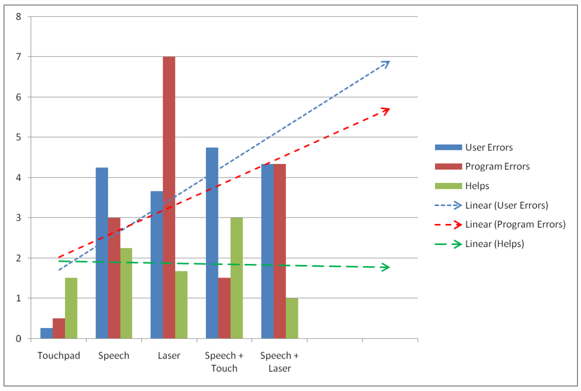


Figure 3. Errors per modality.

Command/Task Analysis. The minimum commands to correctly trigger the task increases as the number of the task rises. However, each task was created to evaluate particular aspects of the interaction. Tasks that used more diagonal movements showed a 280% increase in user errors. Tasks that required higher navigation commands cardinality had more program errors.

4.2. Performance Evaluation

The users performance was captured while interacting with the system by the completion times of each task. We expected similar times in all users but their particularities created some heterogeneity in the results (Fig. 4). Results varied from 28.82 to 44.34 seconds in average, being the laser pointer the input with the worst results and the touchpad with the better results. The most constant user, NC, had times from 35-45 seconds in every modality. All others, surprisingly achieved better times when using the Touchpad which is a gesture-based device. Speech demonstrated the same behavior showed by the error analysis; the times were reduced by 6.29% when using speech with the touchpad and by 5.05% when using speech with the laser pointer. When comparing the times of the laser pointer with the combination with speech surprising results appeared, a reduction of 15.24%. These results clearly showed a reduction of times when comparing with only one device or modality, except in the case of the touchpad which had great results and the combination with speech increased its times by 28.71%.

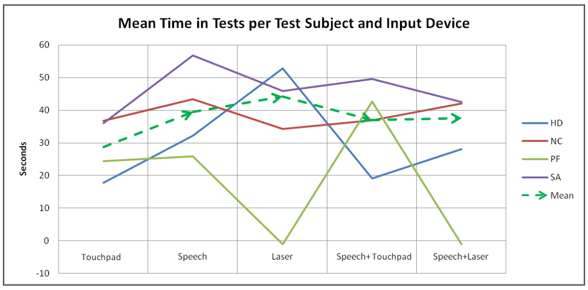


Figure 4. Times taken per modality.

4.3. Usability and User Satisfaction

Through the questionnaires we could assess the users satisfaction with the inputs used and with the system. All users pointed the laser interaction as the less comfortable input method, however they classified it as intermediate in the comfort aspect. The speech was found the easiest and the most comfortable to operate. The mental effort to operate the system was found as residual in all devices or combinations. The physical effort was only an issue when using the laser pointer which, due to its wide movements, was found of medium effort.

4.4. Discussion

The touchpad showed the best results considering performance and error rates with an advantage of at least 35% when compared to all other (multi-)modalities. During the tests, the gesture-based tasks that navigated more in the house hierarchy and the tasks which used more diagonal movements presented more errors than other tasks with a larger number of commands. When asked after the tests users were both impressed and satisfied with the performance of this device. Speech demonstrated high error rates and medium performance. However it was the users preferred method, and its high error rates could be explained by noise and system speech feedback (captured by the microphone creating a loop). Although some problems existed when using speech, its combination with the laser demonstrated a performance increase of 5,05% when compared with only speech. When mixing speech with the touchpad results showed a 6,29% performance increase and 50% program errors, when compared with speech. Clearly, speech interaction, when mixed with other gesture-based devices, could improve users accuracy and performance. Laser pointer had the worst overall results.

5. Conclusion and Future Work

We have presented a system that enables tetraplegic users to control their household appliances. This was achieved by using low-cost control interfaces and interaction methods suitable for the target population. Results showed a high user satisfaction and easy interaction, although further studies with more users are needed to increase the confidence level of these statistics. For the target user population the system has shown great value and has the potentiality of increasing user independence, presenting times for activating 4 devices in a bedroom under 1 minute for all users. Gestures using the touchpad presented the best results either in accuracy and in performance, while the laser pointer presented the worst results which could be explained by its free-hand interaction.

6. Acknowledgments

Tiago Guerreiro was supported by the Portuguese Foundation for Science and Technology, grant SFRH/BD/28110/2006.

References

[Aloise 06] F. Aloise, F. Cincotti, F. Babiloni, MG Marciani, D. Morelli, S. Paolucci, G. Oriolo, A. Cherubini, F. Sciarra, F. Mangiola*et al.* *The ASPICE project: inclusive design for the motor disabled.* ACM Proceedings of the working conference on Advanced visual interfaces, pages 360–363, 2006.

[CEN 03] CEN. *Guidelines to Standardisers of ICT products and services in the CEN ICT domain.* Rapport technique, European Committee for Standardization, 2003.

[Cherubini 07] A. Cherubini, G. Oriolo & F. Macr. *Development of a multimode navigation system for an assistive robotics project.* IEEE International Conference on Robotics and Automation, 2007.

[Do 05] J.H. Do, H. Jang, S.H. Jung, J. Jung & Z. Bien. *Soft remote control system in the intelligent sweet home.* IEEE/RSJ International Conference on IROS Intelligent Robots and Systems, pages 3984–3989, 2005.

[Hill 00] F.J. Hill. Computer graphics using opengl. Prentice Hall PTR Upper Saddle River, NJ, USA, 2000.