An Ergonomic Approach to Computer Vision Interaction

Pere Ponsa, Cristina Manresa-Yee, Marta Díaz, and Cecilio Angulo

***Abstract*—Diverse studies have been made applying human-computer interaction knowledge and techniques, such as accessibility and usability, to help people with severe disabilities carry out specific tasks with a computer. One of the aims of this paper is to show the relationship between assistive technology and human-computer interaction, specifically in the use of vision-based interfaces (VBI). The second part of this paper proposes a framework with the aim of analyzing vision-based interfaces prototypes in order to improve their efficiency. Next, experimental tasks are presented. The first one is an ergonomic guideline approach to VBI. And the second one is an experimental test in order to evaluate the mental workload of the user with the NASA-TLX method. Finally, this paper proposes a task performance method in order to evaluate the accuracy of the VBI and the user experience.**

I. INTRODUCTION

A recent study on the impact and wider potential of information and communication technologies [8] outlines the needs of the growing market of elder people and people with disabilities:
-SMS-based emergency services are useful for deaf people.
-Adaptable and adaptive user interfaces will make systems easier to use by people with disabilities.
-The potential of ambient intelligence to help people with disabilities is considerable.

One of the main difficulties we need to solve is related to the design problem. In the sense that the design is not so concerned with the creation of new technical artifacts as it is with their effectiveness and integration. To some extend this can be reduced to fix the interaction between subsystems such as:
-People with special needs (cerebral palsy, multiple sclerosis, traumatic brain injury) [2], [3]
-Vision-based interfaces (eye-gaze tracking, nose tracking) [4]
-Tasks (home system, computer access, people communication)

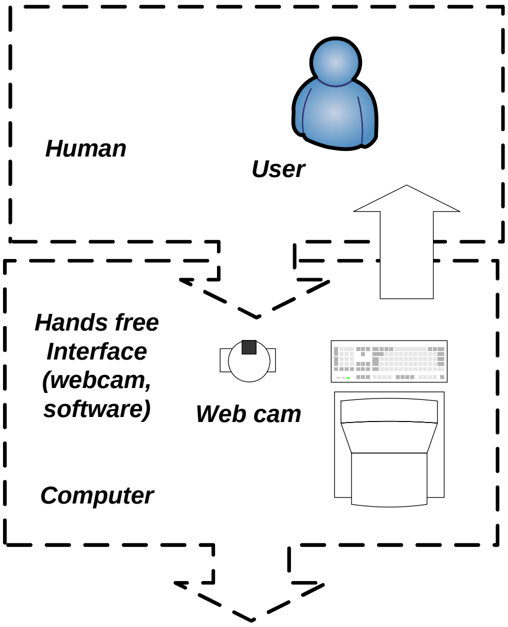

Fig. 1. Human-Computer system. The interface is a hands free interface with a webcam and a specific software.

We are interested in the relationship of these subsystems in order to improve the efficiency of the human-computer system.

If we consider a perspective of user centered design, many disciplines come together. Some of them are: assistive technology, ergonomics, accessibility, ambient intelligent, information technologies, and human-computer interaction.

An important task is the measurement of the usability and the accessibility, carried out in experimental tests with users in a usability laboratory. Moreover, it is necessary to take into account the measurement of the user satisfaction.

The paper is structured as follows. The second section of the paper presents the human-computer system overview, a brief description of the subsystems and the framework of our study. The third section presents two applications. The first one is the study of a vision-based interface VBI prototype; the second one is the use of a commercial vision-based interface inside a home system simulation. In this section, ergonomic guidelines are presented in order to improve the efficiency of the prototype.

Next, this paper shows the use of the NASA-TLX method in order to evaluate the mental workload of the users, a group of engineering students. The fourth section shows a proposed performance method in ordeer to evaluate the usability properties (efficiciency, satisfaction). Finally the conclusions and future steps are presented.

Manuscript submitted April 21, 2008. This work has been partly supported by the Spanish MEC project ADA-EXODUS (DPI2006-15630-C02-01).

P. Ponsa. Automatic Control Department, UPC, Technical University of Catalonia, Av. Víctor Balaguer s/n, 08800 Vilanova i la Geltrú, Barcelona, Spain; phone: 034938967231; fax: 034938967700; e-mail: pedro.ponsa@upc.edu

C. Manresa-Yee. Departament de Matemàtiques i Informàtica. Universitat de les Illes Balears. Ed. Ansel Turneda, Crta Valldemossa km. 7.5 07122, Palma, Spain; (e-mail: cristina.manresa@uib.es).

M. Díaz. 4all-L@b, Usability Laboratory. UPC, Technical University of Catalonia. Av. Víctor Balaguer s/n, 08800 Vilanova i la Geltrú, Barcelona, Spain (e-mail: marta.díaz@upc.edu).

C. Angulo. Automatic control department, UPC, Technical University of Catalonia, Av. Víctor Balaguer s/n, 08800 Vilanova i la Geltrú, Barcelona, Spain; phone; (e-mail: cecilio.angulo@upc.edu)

II. SYSTEM OVERVIEW

The human-computer system has two subsystems, the human and the computer.

One of the main references in telerobotics, automation and human supervisory control is Thomas B. Sheridan, who has influenced in professionals of the areas of engineering of systems and human factors along two decades [5]. One of this author's main ideas is that the progresses in robotics do not depend only in changes in technology, but also in the advances in the understanding of the relationship between people and machines. So, it is necessary to focalize our interest in the interaction between these subsystems (human, computer) using an effective method.

A. *The Human*

The user under the research study is people with or without disabilities. Some authors have made studies with people with severe disabilities, for example cerebral palsy, multiple sclerosis, or traumatic brain injury. In these cases it is very important the impact in the users' quality of life, the families and their caregivers.

B. *The Computer*

Vision-based interfaces VBI use vision computer techniques for human-computer interaction (HCI) purposes and are a kind of Perceptual User Interfaces (PUIs). These interfaces seek to make user interfaces more natural and compelling by taking advantage of the ways in which people naturally interact with each other and with the world [6]. The aim is the development of a no intrusive, comfortable and reliable interface that is easily adaptable.

The importance of VBI grow when the user is physically disabled with limitations in upper body limbs (i.e. cerebral palsy, multiple sclerosis) and can not make use of the traditional input devices such as a mouse or a keyboard. In order to achieve a human-computer interface by means of computer vision and for it to be efficient and suitable, it has to encounter several requirements referring to speed, precision, accuracy and robustness.

This subsystem can help in the e-accessibility component of the e-inclusion by offering a hands-free access to the computer. Moreover it is a low cost subsystem and no cumbersome device is attached to the user making him feel more natural in his communication with a computer.

C. *Method*

An example of methodological framework is the Process Model of the Usability Engineering and the Accessibility MPIu+a developed by Toni Granollers which gathers together all the cycle phases: requirements' analysis, design, implementation, launching, prototyping, evaluation and user) [7].

In the evaluation phase, and for the usability measure, it is necessary to have the contribution of the experimental studies carried out in the usability laboratories.

Table I. Brief analysis of requirements. The factors are: user, VBI and the task

Requirements	**Description**
User	
Age	[child, young, elder]
Gender	[male, female]
Condition	[cerebral palsy, multiple sclerosis, traumatic brain injury]
Vision-Based Interface	
Camera Mouse	Corneal and pupil reflection
EagleEyes	Electrooculographic potencial EOG
Eyetracker	Iris tracking
SINA, Facial Mouse	Nose tracking
Task	
Home system	Shared control of home displays (switches, lights, devices)
Computer Access	Send an e-mail Web navigation
Communication	Family Caregiver Physician

The analysis of requirements is a necessary previous research work in order to establish the best relationship among the human, the interface, and the task. For example, some previous research works are focused in particular cases:

-Example 1: people (cerebral palsy), VBI (Corneal and pupil reflection), Task (Computer Access)

-Example 2: people (cerebral palsy), VBI (Nose tracking), Task (Computer Access)

The problem is that cerebral palsy is a group of chronic conditions affecting body movement and muscle coordination and therefore, it causes disorders in the development of movement and posture, so it isn't correct to group users with different diagnostics in the same experimental task. Another problem is: which is the best interface, from the point of view of the user's experience? Some interfaces are appropriate for a few users but no for all users. It is difficult to obtain generic conclusions, though it's a fact that experiences with vision-based interfaces are very encouraging.

Next, an example of the prototyping phase and an example of the evaluation phase are shown in the third section of this paper.

III. HUMAN-COMPUTER INTERACTION STUDIES

In this section we explain an ergonomic approach over a VBI system and the evaluation of the mental workload of the user with two vision-based interfaces. SINA was used in the first validation and Facial Mouse of CREA

Fig. 2. VBI prototype: SINA

was used in the evaluation of the mental workload.

A. *Testing a visual-based interface prototype*

It corresponds the experts in usability engineering to define how the specifications are evaluated, how to collect the user's opinion and keep it in mind inside the design cycle, and finally as summing up the minimum number of prototypes starting from the one which, the iteration of the cycle is already considered enough to give had concluded the design [8].

The Universitat de les Illes Balears, UIB, has provided us with a vision-based interface prototype called ‘Advanced Natural Interaction System’ (SINA in Catalan language) [9]. This interface incorporates a graphical tool bar that appears on the screen’s right side.

First, it has been necessary to test the possibilities of using the icons of the event selection graphical keyboard of the face-based interface (Fig. 2) in order to work with the blocks of elements on the screen (in order to access a web page, open a Word document, send an e-mail with a specific educational software, etc.).

In this aspect the use easiness and the comfort for the future user has been looked for. Starting with a brief explanation of the functioning of the face-based interface, the user carries out several head movements to check if the application follows the movement of the nose and it allows him the control of the objects on the screen.

Human Factors engineering and cognitive ergonomics can help us in order to validate the characteristics of the prototype. The authors have developed an ‘ergonomic guideline for supervisory control interface design’, the GEDIS guide, like a method that seeks to cover all the aspects of the interface design [10], in order to improve the effective human-computer interaction applied to supervision tasks.

In order to improve the efficiency, before applying the VBI on a user with severe disabilities, it is necessary to validate the prototype. Some factors to study are: the calibration of the VBI, the user-oriented graphical tool bar, the head motion and the feedback. Table II show the ergonomic recommendations.

Table II. Ergonomic recommendations for the VBI prototype SINA

VBI Prototype: SINA	**Ergonomic recommendations**
Calibration/Recalibration	
The calibration process is too difficult to understand for very young children A heavy calibration can influence in the user’ satisfaction	Reduce the calibration process at minimum in order to obtain a natural interface Training the user with useful tasks or games
Graphical Tool Bar	
A graphical tool bar is more intuitive that a text tool bar The navigation inside a text tool bar can be difficult for some users	Improve the location and visibility of the graphical tool bar The tool bar must be easy to understand and use
Head Movements	
The repetitive head movements of the user can increase the fatigue Some users have reduced head mobility	It is necessary a correct relationship between the head movement of the user and the pointer movement on screen It is necessary to guarantee a good performance with a low number of head movements
Feedback	
The user action and the consequence of the action are linked in a correct interaction The user must have the total control of the software application	Quickly transition between the action and the consequence of the action (fast speed and robustness of the tracking algorithm) In some cases itwill be necessary the aid of the caregiver or an artificial software assistant

B. *Evaluation of the mental workload of the user*

Often new technologies are introduced to relieve some very high cognitive demands of the people with severe disabilities. However an automation excess can produce the opposite effect and increase the workload.

The introduction of new interfaces should not increase too much the user's mental workload; a contradiction would take place with the aim to help people with disabilities with a new tool that surpasses his capabilities.

The experimental study was done in June 2007. Inside a laboratory of the Technical University of Catalonia with the Facial Mouse of CREA enterprise, a QuickCam Connect of Logitech, a computer PC (Windows XP, Pentium processor 1500 Mhz, RAM 504 MB), and a display created with Intouch software of Wonderware

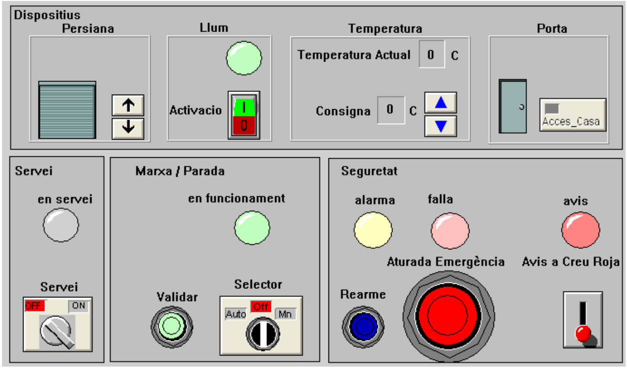

Fig. 3. Home system display.

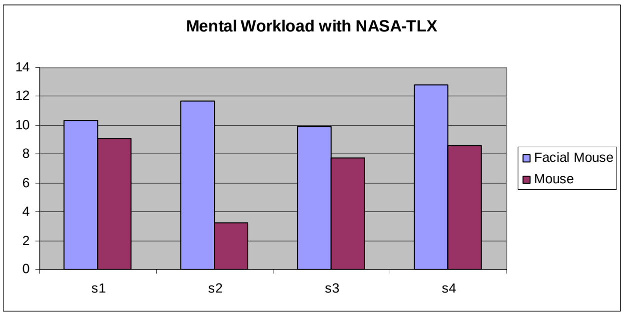

Fig. 4. NASA-TLX task load index method. Where s1 is the user 1, etc.

enterprise. The participants were a set of four electronic engineering students. The experimental methodology was a qualitative evaluation in:

-Comparison between the use of the mouse versus the use of the Facial Mouse

-Location of the webcam: over the screen (left, center, right), over the table (near the computer)

-Use of a display in a home environment

The display of Fig. 3 has two parts. The bottom level is a set of switches and the visual information displays. An example of an important device is the A/M switch; the issue of this device is the transition from automatic to manual control (*human in the loop*). The top level is an example of devices in home systems: lower/raise the blind, turn on/off the light, change the set point of the temperature loop, and open the outside door. Finally, in bottom right section, the user can send an external signal to the Spanish medical service.

Each participant had two trials sessions, the first one with the mouse, and the second with the Facial Mouse. The duration of each session was variable, from 1 to 3 minutes. The instructor gave instructions in order to lower/raise the blind, switch on/off the light, change the set point of the temperature loop, and open the outside door. When the users ended the practice phase, the instructor used the NASA-TLX questionnaire to evaluate the mental workload.

The NASA-TLX task load index method is used for subjective assessment of user workload. The relevant factors for workload assessment are: mental demand, physical demand, temporal demand, performance, frustration level and effort. The user gives a numerical value for each factor and the weight of each factor. Finally, it is possible to obtain the task load index. We show in the Fig. 4 that the facial mouse demands more mental workload than normal mouse, but in this case the results would be also influenced by the experience of the users with the normal mouse and not with the facial mouse.

IV. HUMAN PERFORMANCE AND VBI PERFORMANCE

In human-computer studies, in example the assistive technology applied to people with severe disabilities, it is necessary to define qualitative and quantitative performance rates. Following some ideas of the experts in usability studies and field studies, the tasks presented in the previous section demand a high level of activity planning that involve reasoning and decision making. It is possible to follow different approaches: the individual differences approach, the case study approach and the system characteristics approach. In this paper we follow the first approach. The studies of user's differences have diverse goals:

- To find ways of predicting performance
- To find and characterize individual variability. To find not only differences in the degree to which users are able to reach the goals, but also differences in how they perform, i.e. decision making strategies and user satisfaction

The proposed performance method can summarize in four steps.

- Step I: Accuracy and repeatability test
- Step II: Fitt's law adaptation
- Step III: Miller's magical number adaptation
- Step IV: Task performance index

Step I: to evaluate the accuracy and repeatability of a VBI prototype, the Universitat de les Illes Balears tested a set of 22 users, where half of them had never experienced with the application and the other half had previously trained for a short period with the VBI prototype. A 5x5 point grid was presented in the computer screen where the user had to try clicking on every point; each point had a radius of 15 pixels. While the user performed the test task, distance data between the mouse's position click and the nearest point in the grid was stored to study the accuracy. The error distance is the distance in pixels of the faulty clicked positions (clicks that weren't performed on the targets). The performance evaluation results are:

- For the trained subgroup, the recognized clicks were 97,3% and the mean distance of errors was 2 pixels
- For the novel subgroup, the recognized clicks were 85,9% and the mean distance of errors was 5 pixels.

Step II: in the use of a VBI, a typical task is positioning the pointer and making an action over a device. In order to link the Step I with the Step II it's necessary the Fitts's law adaptation.

The Movement Time, MT, is the time taken to carry out the movement from the origin to the target

$$MT = a + bID_e \quad (1)$$

where ID_e is the index of difficult of the task (in bytes) and *a* and *b* are empirically determined constants, in example *a*= 0,2 and *b*=0,125.

$$ID_e = \log_2\left(\frac{D}{W_e} + 1\right) \quad (2)$$

where D refers to the distance to the target and W_e to the effective size of the target.

A performance rate that makes a relationship between speed and precision is

$$TP = \frac{\overline{ID_e}}{\overline{MT}} \quad (3)$$

where $\overline{ID_e}$ and $\overline{MT}$ are mean values of ID_e and MT in an experiment.

Step I shows the minimal size of an object inside the screen. In this Step, we propose that the size of a device on the screen has to be double the value of 15 pixels. In a theoretical example the user makes two movements:

1) from origin to a target with an object of size W_1
2) from origin to a target with an object of size W_2

In this example, the distance D is constant in both cases, but in the first case the objects' size is double the size of the second case.

The Fitt's law shows that the movement time MT is inferior in the first case. Then, the recommendation in the device design on screen is to increase the size of the target-object in order to facilitate the success of the task and reduce the movement time.

Step III: working memory is generally considered to have limited capacity. The earliest quantification of the capacity limit associated with short-term memory was the magical number seven introduced by Miller, in fact this number is seven plus or minus two (the mininum value is 5 and the maximum is 9).

Miller's research shows some limits on our capacity for processing information. In order to link the Step II with the Step III it's necessary the Miller's magical number adaptation. The problem is: how many graphical objects will be embedded on screen?

According to the previous Steps, the 800x600 screen resolution carries out the Miller's magic number. The screen shown in Fig. 5 shows the recommendation of the maximal number of objects on screen: five objects:

- 5 objects : light temperature, blind, television and lamp (*Llums*, *Temperatura*, *Persiana*, *Televisor* and *Lampara* in Catalan language)
- 1 special object: the VBI graphical tool bar (right)

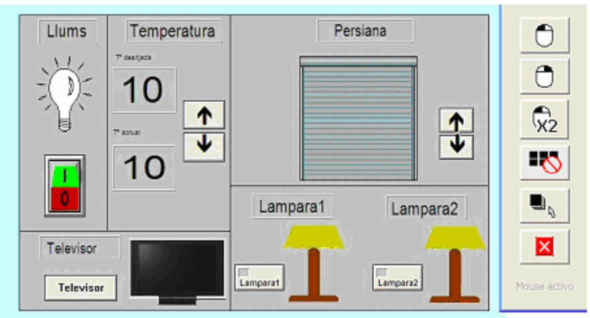

Fig. 5. An example of Miller magical number with dining room and graphical tool bar display

Step IV: Task performance. In simple tasks it is possible to study diverse conditions:

- the user's simple response to an stimulus (the computer can switch on the TV and the user must make the action of switching it off)
- the user must positioning the pointer over a device, in example turn on the light, in order to study the users actions on time domain
- the user's action without external stimulus (the user listens an instruction and then decides or not make an action of a set of actions), in order to study control strategies (or behavioral patterns)

In complex tasks the usability expert can prepare a set of experimental tasks with increasing complexity. For example, the expert prepares an instruction and the user must navigate across screens and finally make a set of actions. In this case, it is important to define a task performance index like

$$TP = \frac{\sum_{j=1}^{N} f\left(T_k, S_j\right)}{N} \quad (4)$$

where N is the total number of subtasks, T_j is the transition from screen_k to screen_{k+1}, and S_j is the success rate of the subtask j.

With this task performance index it is possible to study individual strategies, no in the sense to make a user's classification but in order to study behavioral patterns of the users (with or without disability).

Along an experimental session carried out in the usability laboratory in November 2007, a total of eight subjects (control engineering students) participated in the parameter setting phase of the VBI interface software, the use of the display 'dining room'. The average duration value used for the test was of 96 seconds, the minimum length value of the task was 82 seconds and the maximum length value of the task was 103 seconds.

The instruction: from the "home system" screen it is necessary to choose the dining room switch and go to the "dining room" screen. In this screen the user must carry out five actions. When the user has finished, it is

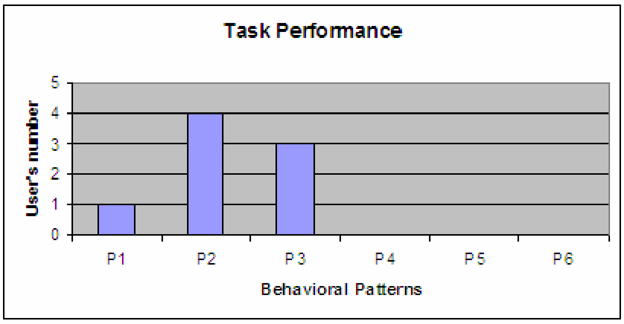

Fig. 6. Task performance in an experimental task with VBI prototype SINA and the 'dining room' interface.

necessary to access to the navigation tool, click on the TV switch and go to "TV" screen. In this screen the user must carry out five actions (choose a TV channel, turn on, turn off, etc.). When the user has finished, it is necessary to come back to the "home system" screen.

In this case:

- N = 2
- T1 from home display screen to dining room screen
- T2 from dining room screen to TV screen
- K = 2
- S1 is the successful rate of the subtask 1. If S1 is equal to five, the user has made a good test. If S1 is 3, for example, the user had made two errors

With the quantitative numeric values of the task performance index it is possible to establish a set of behavioral patterns.

$$\textit{Behavioral Patterns} = [P1, P2, P3, P4, P5, P6]$$

where

P1: the user's performance is correct (the successful rate is the maximum) and the execution time is minimum (82 sec)

P2: the user's performance is correct, and the execution time is regular (82 sec < t <= 96 sec)

P3: the user's performance is correct, and the execution time is high (96 sec < t <= 103 sec)

P4: the user's performance is regular; the successful rate isn't the maximum value

P5: the user's performance is low. The user chooses a wrong path and arrives to another screen (kitchen for example). It is necessary a human error analysis (error, lapse, mistake)

P6: the user makes nothing. The user doesn't navigate across the screen. There is a cognitive problem.

The Fig. 6 shows the number of users that follow the behavioral patterns. The 8 users are distributed in the patterns P1, P2 and P3. From the point of view of a correct performance in a minimum time, one user follows the pattern P1, four users follow the pattern P2 and finally, three users follow the pattern P3.

V. CONCLUSIONS

Computer vision-based interaction is an emerging technology that is becoming more useful, effective, and affordable. In this paper, the authors build the relationship between the use of a visual-based interface with a simulated task: the use of a graphical display in home environment. In future studies is necessary to develop ergonomic validation techniques to reduce the mental workload in the use of visual-based interfaces: in example, is necessary to improve some parameters like the distance of the webcam to the user, the speed of the head's movement in order to minimize the physical demand.

Both vision-based interfaces presented in this paper are currently being tested in centers which work with people with severe disabilities (cerebral palsy). In the case of the vision-based interface prototype called SINA, the main comments of the therapists tutoring the users are related to the interface's usability for improving the user's performance. The most important conclusion offered to us by the therapists is that users (mainly the child) are really excited about accessing to the computer. In addition, users with reduced head mobility have improved its mobility by means of using this interface, proving its potential as a rehabilitation tool.

Vision-based interfaces are a good example of accessible design technology, in the sense that it is possible to extend their use in order to maximize the number of potential customers (child, young people, elderly people).

ACKNOWLEDGMENT

The authors would like thank César Mauri, member of the CREA enterprise, for his collaboration and kindness for allowing the use of the visual-based interface called Facial Mouse.

REFERENCES

[1] R. W. Roe. "Towards an inclusive future". COST219 Project publications, 2007

[2] K. Abe, S. Ohi and M. Ohyama. "An eye-gaze input system using information on eye movement history". 12 th International HCI conference, Beijing, LNCS, Vol 4555, pp. 721-729. Springer, 2007

[3] K. Morimoto, Y. Okuyama, X. Xiaonian, R. Hyun-Seok, K. Kang and S. T. Won. "Strategy of visual search of targets on screen through eye movement of elderly people". 12 th International HCI conference, Beijing, LNCS, Vol 4554, pp. 975-981. Springer, 2007

[4] M. Betke, J. Gips and P. Fleming. "The Camera Mouse: Visual Tracking of Body Features to Provide Computer Access for People with Severe Disabilities". IEEE Trans on neural systems and rehabilitation Engineering 10, 2002

[5] T. B. Sheridan. "Telerobotics, Automation and Human Supervisory Control". M.I.T. Press, 1992

[6] M. Turk and G. Robertson. "Perceptual User Interfaces". Communications of the ACCM 43, 32-34, 2000

[7] C. Mauri, T. Ganollers and M. Garcia. "Computer vision interaction for people with severe movement restrictions". Human Technology Journal, Vol 2(I), pp. 38-53, April, 2006

[8] B. Shneiderman, "Designing the user interface: strategies for effective human-computer interaction". Addison Wesley, 1998

[9] C. Manresa-Yee, J. Varona and F. J. Perales, F.J. "Towards hands-free interfaces based on real-time robust facial gesture recognition". AMDO'06, LNCS 4069: 504-513, 2006

[10] P. Ponsa, and M. Díaz, "Creation of an ergonomic guideline for supervisory control interface design". 12 th International HCI conference, Beijing, LNCS, Vol 4562, pp. 137-146. Springer, 2007