Journal of Physics: Conference Series

Visual communication interface for severe physically disabled patients

To cite this article: M J Savino and E A Fernández 2007 *J. Phys.: Conf. Ser.* **90** 012093

16th Argentine Bioengineering Congress and the 5th Conference of Clinical Engineering IOP Publishing
Journal of Physics: Conference Series **90** (2007) 012093 doi:10.1088/1742-6596/90/1/012093

VISUAL COMMUNICATION INTERFACE FOR SEVERE PHYSICALLY DISABLED PATIENTS

M J Savino[1], **E A Fernández**[1, 2]

[1] Fac. de Ingeniería, Universidad Católica de Córdoba, Camino a Alta Gracia km. 10, (5000) Córdoba, Córdoba, Argentina

[2] *CONICET*

maurisavino@gmail.com

Abstract. During the last years several interfaces have been developed to allow communication to those patients suffering serious physical disabilities. In this work, a computer based communication interface is presented. It was designed to allow communication to those patients that cannot use neither their hands nor their voice but they can do it through their eyes. The system monitors the eyes movements by means of a webcam. Then, by means of an Artificial Neural Network, the system allows the identification of specified position on the screen through the identification of the eyes positions. This way the user can control a virtual keyboard on a screen that allows him to write and browse the system and enables him to send e-mails, SMS, activate video/music programs and control environmental devices. A patient was simulated to evaluate the versatility of the system. Its operation was satisfactory and it allowed the evaluation of the system potential. The development of this system requires low cost elements that are easily found in the market.

1. Introduction

Human physical-motor disability is a condition in patients with some prenatal problems or that have acquired this condition in an accident or that have had problems in the first years of their life [1].

The number of physically disabled people in the world has increased as a result of many different causes. Many of these are related to traffic accidents where, in most cases, the consequences are very severe because they affect not only physical-motor abilities but also intellectual ones.

In many cases the patient realizes about their disabilities. Here the communication possibilities are crucial in order to improve the quality of life of these patients. However, in many cases the patients are unable to speak due to traumatic experience causing their actual situation.

There are many types of disabilities depending on diverse factors such as intensity, complexity and durability of this condition. Because of this, it is pertinent to acknowledge that the needs from one user to another are different and this can result in the development of a completely different final system used as a communication interface. Nevertheless, a system should have a certain degree of simplicity and should be as user-friendly as possible because we might be working with a physically disabled person with diminished intellectual capacities as a consequence of the trauma.

The needs of interfaces allowing common ways of communication, as a computer, are continuously demanded in order to provide the patients with a way to express their feelings and their necessities. These solutions improve their life quality.

Different methodologies exist to tackle the communication problem of these patients. Many of them use goggles on which they mount a video camera that takes images from their eyes [2,3]. There are other more complex systems such as the electronic sensors that catch the sensorial stimuli of the brain [4], being some of them invasive.

Although there are many methods developed in the market, this system differentiates from others by its simplicity in use since it displays very simple interfaces. It is also more affordable compared to other systems because it only needs a simple webcam. In addition, it is highly adaptable to the patient requirements.

In the present work the methodology that was approached is the use of a simple webcam that does not have to be mounted on the face of the patient. The system was designed to maximize comfort in order to diminished fatigue using this interface.

A communication interface was developed that allows the user to interact with its atmosphere through simple ocular movements.

2. Materials and methods

The goal of this system is to allow severe physically disabled patients to communicate through the movement of their eyes registered by a webcam. With such movements the patients can manage a virtual keyboard and a series of virtual instruments for communication and control.

In order to control the keyboard, four regions were defined on the screen or monitor: above, down, left and right and an "accepted" state. The four regions are shown in figure 1. The "accepted" state is defined with a blink of the eyes.

In order to be able to determine the position of the eyes, images of the face are registered with a webcam at a predefined sampling frequency (adjusted by the therapist). Then, by means of open source face detection algorithm [8] the eyes are detected as well as five characteristic points associated with them: the outer and inner points of the eye, their center, the direction of the pupil and the center of the eyebrow (see Figure 2).

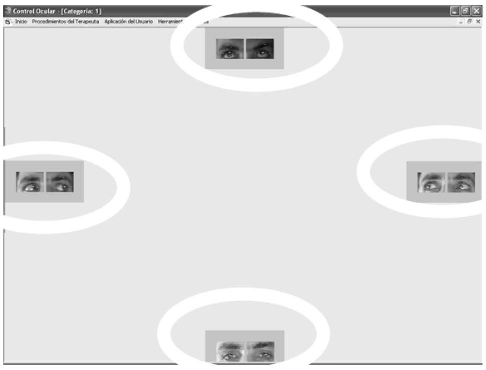


Figure 1. Different regions of control for the obtaining of samples.

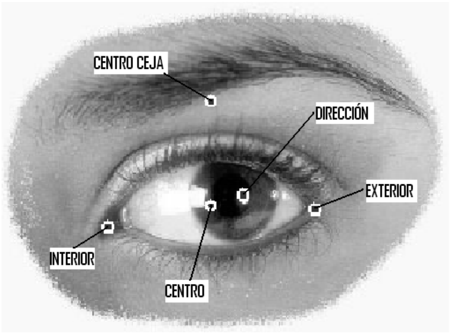


Figure 2. Eyes characteristics

In order to classify the obtained samples an Artificial Neural Network (ANN) was used [6,7]. The ANN determines in which of the five classes (left, right, upper, bottom and accepted) the patient is looking at.

Another characteristic of the system is that it is configurable for each patient. This is done by means of the intervention of the therapist, who makes the process of acquisition of the initial images. Then the system automatically chooses the best neural network architecture for this patient.

This process is made in the following way: the therapist selects one of the four areas defined on the monitor and instructs the patient to direct their eyes to that position. After that, the image is recorded and stored in the system. This process is repeated several times for each of the five classes.

The design of the control system of the interface was based on the Knowledge Discovery Methodology [6,7], which describes the different stages for handling, pre-processing and mining the stored data. As stated above, the data mining step was carried out by means of an ANN model, the Multilayer Perceptron [5].

Prior to feed the ANN, the stored samples are normalized. The normalization method is selected between "Zscore" and "Min-Max" [6] from the configuration panel of the system.

In order to choose the best architecture selection, (the appropriate number of neurons in the hiden layer HLN) - the system spans different numbers of HLN and validates each of them by means of a cross-validation technique [6,7]. The range of neurons in hidden layer is settled by the therapist. The best model is automatically chosen by the system based on the minimal classification error on the cross-validation process.

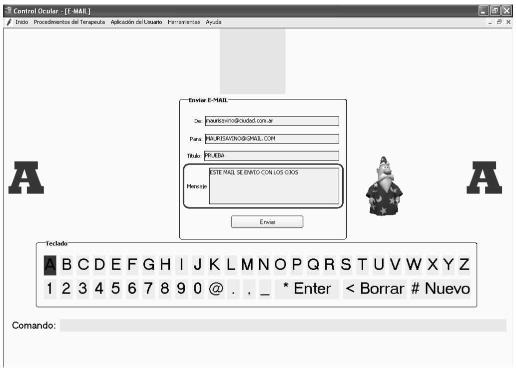
Once selected the suitable model, the therapist leaves the system already working. In this way the patient uses the screen as a control instrument through which he writes commands to be executed. The system takes samples from the position of the eyes, extracts the characteristics and then, classifies them in some of the possible categories. After this classification, the system reacts in a particular way (writing by means of the virtual keyboard) according to the predicted class. By means of this, the user can ask for programmed commands and thus be able to interact and control some actions. These actions let him interact with the therapist and his relatives as well. The actions that can be made in this interface are only examples of what can be done if the position of the eyes is detected.

Some of these system functionalities are the e-mail service, SMS cell phone messages, program execution - audio, video, executable files-, control of parallel port to on/off electrical or electronic devices and the reproduction of the text written through Microsoft agents [9], among others. Figure 3 shows the interface of e-mails with its virtual keyboard, the active key in the lateral ends and the picture that indicates where was the patient looking at (shadowed square on top of the screen). The screen was opened by means of the command "MAIL" and thus the patient is able to write the message on the focused screen area.

The user/patient moves the cursor through the keys fixing his/her glance to the left or right point of the screen. The cursor is moved towards that direction. If the user wants to move the cursor more than one position, he fixes the glance towards the above region. By doing this, the user can navigate through the different controls.

Once the user is positioned over the desired virtual key or function, the selection is done by blinking his/her eyes during a certain period of time (“accepted” class). The system gives visual (by writing the chosen letter) and audio warnings.

A simulation environment is presented in Figure 4. One hundred samples from each class were recorded to overcome the test.



De: maurisavino@ciudad.com.ar
Para: MAURISAVINO@GMAIL.COM
Título: PRUEBA
Mensaje: ESTE MAIL SE ENVIO CON LOS OJOS

Figure 3. E-mail command window

Figura 4. Visual interface and its components

3. Results

In Table 1, the results for the different cross-validated ANN with different numbers of HLN are shown. The average classification rate is presented as well.

The HLN range was specified from 2 to 25. The best architecture was the ANN having 21 HLN and it had an average training percentage classification rate of 97.11%. The ANN showed a final evaluation percentage near to 89%.With this architecture the simulated patient could conduct all the tasks available in the system.

To avoid confusion from involuntary movements of eyes and blinks, the system expects N successive identical classification over the same class where N≥2 is settled by the therapist in the configuration panel of the software. After this N identical classification the system performs the desired action.

TABLE I

FINAL RESULTS OF VALIDATION

(AVERAGE OVER 6 PARTITIONS OF TRANNING AND VALIDATION).

HLN	Validation Mean % classification rate
2	71,55
3	95,33
4	92,88
5	95,11
6	93,55
7	94,88
8	93,33
9	96,22
10	95,11
11	95,11
12	95,11
13	96
14	94
15	94,66
16	95,33
17	94,66
18	94,66
19	94,66
20	95,33
21	**97,11**
22	94,66
23	94,66
24	95,33
25	94,88

4. Discussion

A prototype was presented here and demonstrated to be versatile in the evaluated simulations. The advantages of this system in patients with severe motor difficulties are very important because they provide independence and communication for those patients with friends, relatives and their therapist. The system provides a high level of automatism that diminishes the intervention of the therapist. This intervention only takes place in the configuration stage and in the initial learning phase of the ANN. After that, the system can be used by the patients on their own.

The use of the system allowed the simulated patient to communicate efficiently with the environment.

In addition, the system is not invasive, and the user should not wear any device on his face or body. It just needs a webcam fixed to the monitor.

Once the therapist takes the samples from the patient eyes, it runs automatically, choosing the appropriate ANN solution for this specific patient.

The system is versatile enough to be adjusted to specific patient requirements in an easy way by means of different configuration settings and by means of the specific trained ANN.

An overall technical solution does not exist for this kind of patients. A more complete solution includes the use of technology and the participation of their relatives and therapists, among others. Joining all these factors together it would be possible to offer an integral solution in order to improve the life quality of such patients.

References

[1] "La situación de los discapacitados en la Argentina"
http://www.discapacitados.org.ar/documentos/documento1.doc
[2] "Interfaz para Control de Computadora con el Movimiento Ocular" (www.keyes.com.ar)
[3] "Sistema EyeLink" (http://www.eyelinkinfo.com)
[4] "Interfaz Cerebro-Computadora" (www.iti.upv.es/cas/papers/CAS-BCI2004.pdf)
[5] Haykin S. Neural networks—a comprehensive foundation. 2nd ed. Upper Saddle River, New York: Prentice Hall. Intern Inc; 1999.
[6] Fernandez, Elmer A. – "Minería de datos y Redes Neuronales para el análisis de la cinética de diálisis" – Tesis de doctorado, 2003
[7] Kennedy, R.L., Lee, Y, Van-Roy, B., Reed, D.C., Lippman, R.P. "Solving Data Mining Problems Through Pattern Recognition." Prentice Hall – 1998
[8] http://sluggish.uni.cc/rodney/components.htm#FaceDetection
http://www.microsoft.com/msagent/default.asp