Learning and Nonlinear Models (L&NLM) – Journal of the Brazilian Neural Network Society, Vol. 9, Iss.4, pp. 249-255, 2011.

AN INTELLIGENT HUMAN-MACHINE INTERFACE BASED ON EYE TRACKING TO AFFORD WRITTEN COMMUNICATION OF LOCKED-IN SYNDROME PATIENTS

Amanda Leonel, Fernando Buarque de L. Neto, Sérgio Campello Oliveira, Hugo Serrano B. Filho

Computer Engineering Programme, Polytechnic School of Pernambuco, University of Pernambuco
Rua Benfica, 455 – Madalena, 50.720-001 – Recife, Pernambuco – Brasil
aln@dsc.upe.br,{fbln, scampello, hsbf}@ecomp.poli.br

Abstract – Patients with 'Locked-in Syndrome' (LIS) report rare and severe clinical cases as they become paralyzed and voiceless. Due to a brain damage, patients lose movements of their voluntary muscles and also,their ability to speak. In spite of that, their cognitive functions remain in perfect order. In general, the only voluntary movement that still remains is the motor control of the eyes and eyelids. This work presents a proof-of-concept,including an algorithm, based on an IntelligentHuman-Computer Interface to enable written communication of paralytic patients. The proposed system uses techniques from Computer Vision and Artificial Intelligence in order to acquire and detect eye-eyelid movements affordingwritten communications of LIS patients. The processed images result in low resolution patterns of the patients'eye-eyelid. Artificial Neural Networks are then trained in order to recognize such patterns and relate them to written symbols that are presented in a computer screen. The high rate of successand performance on the system, obtained in experiments demonstrate the feasibility of the proposed interface model.

Keywords – Human-Computer Interface, Paralysis, Locked-in Syndrome, Eye-tracking, Artificial Neural Networks.

1. Introduction

User interface is a group of resources by which users interact with a system. Human-Machine Interface (HMI) or Human-Computer Interface (HCI) arethe terms used in the context of Engineering and Computing [1]. In order to improve the performance of HCIs, computing techniques as Pattern Recognition, Artificial Intelligence and Software Engineering are often applied in cutting edge HCI.

One possible use of advanced HCI integrated in computer systemscould be to help patients with limited or compromised communication asin the case of patients with Locked-in Syndrome [2]. Locked-in Syndrome (LIS) is a rare condition in which a patient has his/hervoluntary muscles paralyzed (due to a brain stroke, for example). Even though, LIS patients still maintain all their cognitive functions in perfect condition (they can see, hear and understand everything around them), but they cannot interact.

Thus, computer systems equipped with advanced HCIs could make a huge difference for LIS patients.The problem is that these kinds of HCIs are very expensive and theiruse in the rehabilitation on many cases is difficult because operational limitations in acquiring the patients idiosyncrasies.

This work aims to model and implement an intelligent HCI for written communication, based on eye-tracking suitable for LIS patients. The proposed HCI uses Artificial Neural Networks and a low cost computational system. The expected result is to build a proof of concept for modeling and implementation of such HCI that enables communication of Locked-in Syndrome patients as well as presenting some individual customization.

This paper is organized as follows. Section 2 presents the main concepts necessary for the theoretical understanding of the challenges of the selected task. The proposed model is presented in section 3. Insection 4 the preliminary experimental results are presented and discussed. Finally, insection 5 conclusions and suggestions for future works are commented upon.

2. Background

2.1 Locked-in Syndrome

According to World Health Organization (WHO), a syndrome is "a group of signs and symptoms based on their frequent co-occurrence, which may suggest a basic pathogenesis, course, familial pattern or common treatment" [3]. Thus a syndrome characterizes not necessarily as a single disease, but as a group of similar diseases in some aspect.

Locked-in Syndrome was defined in 1966, but since 1986 has been described as "any quadriplegia and anarthria with preservation of consciousness" [2]. *Quadriplegia* is the case because the patient has paralysis of almost all voluntary muscles of the body. In classic cases of LIS, the patient maintains only the eye movements; *Anarthria* is due to their inability to articulate words;and *Consciousness preserved* because the patient retains all cognitive functions, i.e. perception (e.g. vision and hearing), attention, intellectual abilities, verbal and visual memory etc.

Acute ventral pons lesions are the most common cause of LIS. Pons is a structure of the brainstem located between the brain and spinal cord in the human central nervous system and the pons is responsible for the transmission nervous impulses from spinal cord to the brain [4].

In particular, the case of Jean-Dominique Bauby (aged 43 and editor in chief of the fashion magazine *Elle*) made the syndrome widely known. Bauby wrote an unusual biography [5] in which he composed the chapters mentally. A secretary recited a frequency-ordered alphabet until he "dictated" a letter by blinking his one-functioning eye lid.

2.2 The Human Eye

In patients with classical LIS, eye movements are the only one possible voluntary movement. The human eye is composed of several structuressuch as pupil, iris, cornea, lens, retina and optic nerve [6]. The eyeball performs movements around its center. Four main types of human eye movements are [7]:

- Saccadic movements (or *saccades*): are sudden and quick. In reading, for example, saccades generally occurs when looking from one word to next;
- Smooth movements: keep the line of sight on the selected object during the intervals between saccades;
- Fixations: maintain stationary visual stimuli;
- Blink: takes an average of 0.3 to 0.4 seconds. During the reading, an adult can blink only 3 to 4 times per minute.

2.3 Digital Image Processing

Eye movements can be captured by cameras, which generate videos (or series of images) in several formats and rates. A digital image is the materialization of input and the basis upon which many techniques of different areasare applied in this work, namely, areas of Digital Image Processing (DIP), Computer Vision and Image Segmentation [8].

In the matrix-based image representation, each point is called a picture element or pixel. A pixel has a color (value) associated toit. The spatial resolution, i.e. the number of rows and columns of the matrix, represents the dimensions of an image. Filtering operations of images are important because they can provide changes on theircharacteristics inorder toimprove quality or, to obtain“new” information.

2.4 Artificial Neural Networks

Artificial Neural Networks (ANNs) [9] can be applied to problems in areas such as Computer Vision, allowing the recognition and classification of patterns embeddedin images. ANNs are a set of meta-heuristics, i.e. a Computational Intelligence technique biologically inspired and of multidisciplinary use. In ANNs, a set of simple processing units - i.e. neurons - are linked within a network, which can exhibit nonlinear processing skills. Thus ANNs have the capability to solve complex computational problems.

There are countless models, architectures and learning algorithms for ANNs, but Perceptron and Multilayer Perceptron (MLP) networks, trained with the Backpropagation algorithm, are the most used models [9]. Design decisions of an ANN are under the responsibility of the network designer and it depends on the problem to be solved. However, some steps must be followed carefully: (i) data acquisition, (ii) network configuration, (iii) training, (iv) validation and (v) test.

2.5 Related Works

There are many commercial HCIs applications with features that do not require user physical contact forinteraction. However, only few of them can be used by patients with LIS due to their severe interaction limitations. In general, such kind of equipment is expensive, and the requirements for accessibility, usability and security demanded by the patients are very specific and, sometimes, not covered.

The eye-movement tracking has been the most promising approach in cases of classical LIS, because it is possible to be developed at low cost [10]. Along these lines, some studies have been developed using conventional algorithms and classical techniques of Computer Vision, or alternative methods of Computational Intelligence (such as ANNs), but systems that possess written communication skills are not that many[11].

3. The proposed model

The proposed model is based on eye-tracking through techniques of Digital Image Processing and Computational Intelligence, to provide written communication to patients with LIS using personal computers. An overview of the system and its modules are shown in Figure 1 and described in along the following sub-sections.

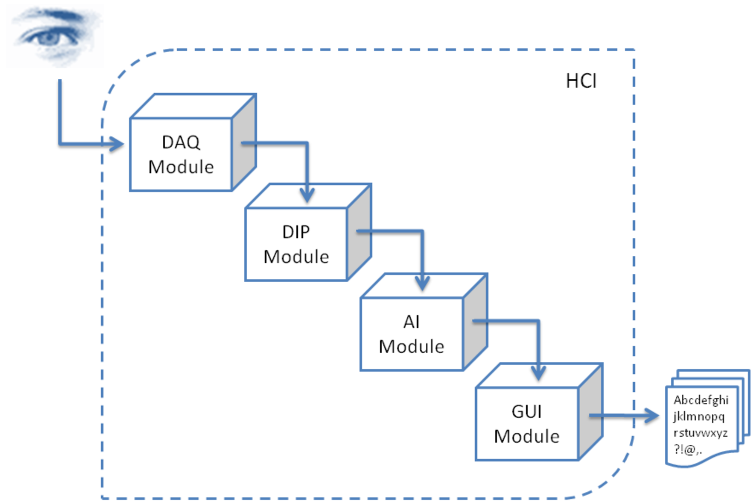

Figure1.General structure and modular organization of the proposed HCI.

3.1 DAQ Module: Data Acquisition

The Data Acquisition (DAQ) module aims to gather input data for processing in a system. In general, it has appropriate sensors to monitor the signals to be acquired.

In this model, the input signals are composed by images of eye movements. The eye-tracking device is based on a webcam with infrared Lighting Emitting Diodes (LEDs) that includes illumination, developed by T. F. Vieira and E. Fontana [10]. It was used to generate the image database required for the proposed system proof of concept.

3.2 DIP Module: Digital Image Processing

The DIP module use classical image processing techniques in order to extract the image data needed for pattern recognition in the Artificial Intelligence (AI) module.

The tool used for image processing was Matlab® version 7 (R2006a) and its complement "Image Processing Toolbox" version 5.2 (R2006a).

The internal processing of the DIP module is ordered as follows:

i. Extraction and classification of the frames;
ii. Image Preprocessing: rotation, spatial resolution reduction and conversion to grayscale;
iii. Pupil segmentation: edge detection, binarization and dilation;
iv. Reduction of image resolution density;
v. Binary conversion.

3.3 AI Module: ArtificialIntelligence

This module defines the architecture and parameters used in the Artificial Neural Network for recognition of directional patterns to be analyzed. The Multilayer Perceptron architecture was the ANN model chosen due to its known ability as a good classifier.

WEKA® version 3 is a software package equipped with Machine Learning algorithms for extending the Java programming language. The module Experimenter from the tool WEKA® was used to perform configuration, training, validation and testing of the MLP.

3.4 GUI Module: Graphical User Interface

This module provides a Graphical User Interface for mapping visual patterns (identified by previous modules) in characters for the writing task. Generically, the GUI module used can be described as a 4x7 grid with 1024x768 pixels of resolution. It contains 28 square fields and a blank rectangle on top, as shown in Figure 2. The squares define the classes used for mapping eye movements and contain letters from the alphabet and other characters for selection when LIS patients will be performing their writing tasks. The rectangular top field is reserved for displaying the set of characters selected at the end of each HCI processing cycle, shown in Figure 1.

1x1	1x2	1x3	1x4	1x5	1x6	1x7
2x1	2x2	2x3	2x4	2x5	2x6	2x7
3x1	3x2	3x3	3x4	3x5	3x6	3x7
4x1	4x2	4x3	4x4	4x5	4x6	4x7

Figure 2. Arrangement of classes in the Graphical User Interface.

4. Experiments and Results

The experiments and results of each module are described in next sections.

4.1 GUI module results

The Graphical User Interface was developed in the Java programming language, to (i) visually map classes (i.e. squares with letters and symbols) into selected characters and (ii) to display the output of the AI module as a text above the screen.

The mapping grid is illustrated in Figure 3. The character ‘<’ is the backspace operation on a normal keyboard. The blank square is a blank space. The ‘@’symbol was added after a report of a patient requesting to include it, so he would be able to write e-mails.

As a result of the GUI module, a small communication protocol was developed for the use of the system. The user must confirm the desired character aiming on its square and blinking for about 1 second. The video capture rate is 10 frames per second, for example. Thus the successive detection of more than 10 images is classified as "blink" and selects for writing the class of the signal of the last image which was not classified as "blink". This approach reduces accidental writing errors, resulting from the human eye natural movement.

A	B	C	D	E	F	G
H	I	J	L	M	N	O
P	Q	R	S	T	U	V
X	Z	@	?	<	.	

Figure 3. Layout of characters and symbols in the graphic interface.

4.2 DAQ module results

The video capture procedure works as follows:

- The device selected and in use was attached to the head of the user. The head was held still (like if it were of a patient with classical LIS), at distance of approximately 30 cm from the computer screen (and camera).
- We asked the user to observe each square of the grid, slowly and successively, going line by line until the end of the grid.
- In each square, the user eyes made smooth movements searching the inner area of the square. These movements allow greater variety of data for the ANN.
- Between square, the eyes made saccadic movements and/or blinks. The blinking indicates explicitly the exchange square.

Finally, the videos captured have the following characteristics shown in Table 1.

Table 1.Technical characteristics of the videos generated.

PARAMETERS	VALUES
File format	.avi
Video compression	no
Number of frames/second	30
Image type	truecolor(RGB)
Vertical resolution (height)	480 pixels
Horizontal resolution (width)	680 pixels

4.3 PID module results

The extraction and classification of frames from the videos generate the 'Alphabet' image database, whose technical characteristics are listed in Table 2. In addition to the 28 classes of the graphical user interface, two extra classes have also been created: 'blink' and 'long saccades'. The long saccadic class contains images of long switches as line breaks. Figure 4 shows samples of the image database generated. The glasses of the DAQ device may generate reflexes in the images.

Table 2.Technical characteristics of the 'Alphabet' image database.

PARAMETERS	VALUES
Image format	.bmp
Image compression	no
Image type	truecolor (RGB)
Vertical resolution (height)	480 pixels
Horizontal resolution (width)	680 pixels
Number of classes	30
Total number of images (aprox.)	5,040
Average number of images per class	168

(a)

(b)

Figure 4. Samples of the original images from different classes of the 'Alphabet' database. (a): Class 'A'. (b): 'Blink' class.

After the image-processing phases ii-iv (see section 3.2), the final images obtained are binary coded (i.e. pixel 1 to white, pixel 0 to black) and have spatial resolution of 13x9 pixels. Final samples from different classes are illustrated in Figure 5.

Figure 5. Samples of the final images from different classes in PID module.

The last image processing phase (v) resulted in the conversion of the final images in ('alphabet.arff') binary database. This binary database was used for training, validation and testing of the AI module in the tool WEKA®. Each pixel of the image represents an attribute in the binary database. The last database of the PID module is described in Table 3.

Table 3. Technical characteristics of the `'alphabet.arff'` binary database.

PARAMETERS	VALUES
Base name (relation)	alphabet
Base format	.arff
Number of attributes	117 (13 x 9) pixels
Number of classes	29 (4x7 + blink)
Number of instances	300 images
Sorting of instances	no (random)

4.4 AI module results

Input (images) and output (classes) desired were presented to the ANN during training in order to adjust the weights and to find a suitable relationship between input-output pairs and for generalization over new data.

The selection of the architecture and parameters of the MLP used was *ad hoc* based on the acceptable results of the training, validation and testing of MLPs studied. The best parameter settings obtained for the ANN, using the 'alphabet.arff' database, is listed in Table 4.

Table 4.Best parameter settings obtained for the ANN.

PARAMETERS	VALUES
Database to:	
• Training	230
• Validation	40
• Test	30
Number ofinput neurons	117 (pixels)
Number of output neurons	29 (classes)
Number of hidden neurons	29 (prop. to output)
Learning rate	0.3
Momentum	0.2
Training epochs	60 (maximum)
Validation threshold	10
Number of repetitions	30 (times)

Table 5 illustrates the results obtained from 30 simulations. This network gets right, on an average, 29.13 images of each set out of 30 patterns presented in the tests, i.e. 96.13% of accuracy. Training takes only 6.46 s approximately.

Table 5.Test results from the best MLP configured.

	Average	Standard deviation	Median
Absolute accuracy	29.13	1.06	29.00
Percent accuracy	96.13	3.53	96.66
Absolute error	1.17	1.06	0.00
Percent error	3.86	3.53	0.00
Mean Absolute Error (MAE)	0.02	0.00	0.02
Root Mean Square Error (RMSE)	0.06	0.00	0.06
Relative Absolute Error (RAE)	30.28	2.97	30.09
Root Relative Squared Error (RRSE)	36.35	4.80	36.12
Training time (s)	6.46	3.71	4.73

5. Conclusions

This work presented and validated a successful proof of concept, proposing the implementation of an intelligent Human-Computer Interface for written communication of patients with Locked-in Syndrome.

HCIs based on eye-tracking can be an alternative communication way for many cases of classical LIS. The special requirements of these patients may facilitate (e.g. lack of head motion) the design of such systems.

In particular, Computer Vision and Artificial Intelligence techniques can ensure robust HCIs, even at low costs. The combination of these techniques does not necessarily impose needs to include state of the art algorithms to ensure satisfactory results. The final images obtained showed representative patterns even in low spatial resolution, eliminating the use of high-resolution images or three-dimensional coordinates in relation to the eyeball.

The use of Artificial Neural Networks was of great importance forthe correct mapping between input and output. In particular, the recognition of different patterns obtained, afforded flexibility and adaptability in human-computer interaction. The accurate of the network simulated shows very good results, close to 100% hits at the final tests. And the short duration of these simulations suggests the possibility of use of such networks in real time, or even at "on-the-fly" training if necessary.

As future work, it is proposed (i) the addition of a real time calibration module in the system; (ii) an automatic parameter configuration of PID and AI modules; (iii) analysis of accuracy of the system, in order to include more classes in the grid; (iv) the execution of actual test; and (v) the implementation of a final product in a programming language compiled, in order to optimize the integration and the computational cost of the system.

With these improvements proposed as future works, an intelligent HIC can act together with an eye capture movement video dispositive to permit LIS patients to communicate in efficient ways with their parents and doctors allowing them for example, to ask for help when necessary.

Special thanks

The authors thank to Prof. Eduardo Fontana and his doctoral student Tiago F. Vieira, from the Photonics Group at UFPE, for their collaboration in this work, enabling the data recording through the device developed by them [10].

References

[1] A. Sutcliffe, *Human-computer interface design*, Springer-Verlag, 1989.

[2] E. Smith and M. Delargy, *Clinical Review: Locked-in Syndrome*, British Medical Journal - BMJ Publishing Group Ltd., 2005.

[3] All and WHO, *DSM-IV: Diagnostic and Statistical Manual of Mental Disorders*, American Psychiatric Association, 1994.

[4] E.R. Kandel, J. Schwartz, and T.M. Jessell, *Principles of Neural Science*, McGraw-Hill Medical, 2000.

[5] J.D.Bauby, *O escafandro e a Borboleta*, São Paulo, Martins Fontes, 1997.

[6] M.J. Hogan, J.A. Alvarado, and J.E. Weddell, *Histology of the human eye: an atlas and textbook*, WB Saunders Company, 1971.

[7] R. Wilson, *Eye Movementes and Visual Attention - The MIT Encyclopedia of the Cognitive Sciences (MITECS)*, The MIT Press, 2001.

[8] J. Gomes and L. Velho, *Computação Gráfica: Imagem*, Rio de Janeiro, RJ: Instituto de Matemática Pura e Aplicada–IMPA, 2002.

[9] S. Haykin, *Neural Networks: a comprehensive foundation*, Prentice Hall, 2008.

[10] T.F. Vieira and E. Fontana, "Dispositivo de rastreamento de movimentos oculares baseado em webcam e iluminação com led infravermelho," *XXI Congresso Brasileiro de Engenharia Biomédica*, Bahia, Salvador: 2008, pp. 669-672.

[11] Metrovision, "Visionboard," 1997, pp. 2-3.Medical and Biological Engineering and Computing, vol. 35, 416, 1997.