

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



---

# Analysis of a Sorter Cascade Applied to Control a Wheelchair

---

Marcos Figueredo , Alexandre Nascimento ,  
Roberto L.S. Monteiro and Marcelo A. Moret

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/63816>

---

## Abstract

The precise eye state detection is a fundamental stage for various activities that require human-machine interaction (HMI). This chapter presents an analysis of the implementation of a system for navigating a wheelchair with automation (CRA), based on facial expressions, especially eyes closed using a Haar cascade classifier (HCC). Aimed at people with locomotor disability of the upper and lower limbs, the state detection was based on two steps: the capture of the image, which concentrates on the detection actions and image optimization; actions of the chair, which interprets the data capture and sends the action to the chair. The results showed that the model has excellent accuracy in identification with robust performance in recognizing eyes closed, bypassing well occlusion issues and lighting with about 98% accuracy. The application of the model in the simulations opens the implementation and marriage opportunity with the chair sensor universe aiming a safe and efficient navigation to the user.

**Keywords:** wheelchair, classifier cascade, detector eyes closed, active vision, eyes state

---

## 1. Introduction

A number of illnesses and accidents can lead to severe damage in the spinal cord of a patient resulting in the loss of motion in the lower and upper parts. According to [1], 14% of the population has some kind of disability, be it visual, motor, hearing, and others, representing a growth of 7% in recent years.

Among this group, about 4% do not have any kind of movement in the lower and upper limbs. This is due to very serious motor problems, such as hemiplegia of four members and the need

for ventilatory support. Also applies to patients with degenerative diseases of the neuromuscular system, for example, amyotrophic lateral sclerosis (ALS) in which in a progressive manner the person loses his movements, until completely paralyzed, thereby causing death by respiratory failure.

Among the various types of motor disabilities that can affect a person, quadriplegia (motor disability of the four members) and the neuromotor system diseases such as ALS are serious deficiencies, which lead the individual to an almost vegetative state, with difficulties with integration into society as useful as a capable person. However, in most cases such individuals have full brain capacity, and with the necessary physical media, they may participate productively in society. Thus, it is necessary to find ways to develop their personal and professional skills and have a professional activity with human dignity.

Typically, this patient uses a wheelchair to perform tasks such as come and go, always with the aid of a carer or relative. Some of these use couplings that allow the user to their locomotion in their environment and in general are very invasive.

The works of [2–5] indicate the concern of the academic community in realizing technology stocks, low cost, that make the everyday life of these more independent people, and these studies also indicate that these actions prolong the lives of patients and improve their quality of life [6, 7].



**Figure 1.** CRA used as prototype.

Several studies have been conducted by different research groups [8–10] to develop wheelchairs with some kind of intelligence or are simply able to understand voice commands, autonomous locomotion, deviation obstructions, and other functions as outlined in [2]. These models have a high cost and maintenance, having little or no type of embedded technology.

Motivated by this reality began an interdisciplinary nature project that aims to provide a common wheelchair with elements that make it both possible and economically viable the mobility of a patient without movements of the upper and lower limbs. Patients with this degree of disability have only a means of interaction with the machine elements present in one's face, which will be the facial expressions and the device that will allow the mobility automation with wheelchair (CRA), as shown in **Figure 1**.

This navigation is limited in some aspects, as some movements of facial expressions were extremely tiring, such as opening and closing the mouth, and others were too difficult to be captured or required equipment that are still under development, such as the retina of the eye movement. So two expressions have proved less susceptible to failure, namely low intrusiveness and low learning curve. They are the opening and closing of the eyes and the movement of turning one's head to right and left. These expressions were evaluated by 20 individuals simulating the state of a quadriplegic and generated evaluation described in **Table 1**.

Facial expression	Failure (%)	Setting (%)	Intrusiveness (0–10)	Learning (0–10)
Close and open your eyes	2	98	2	2
Head spin	3	97	2	3

**Table 1.** Percentage of success and failure detection of facial expressions. Average score of the difficulty of learning to use facial expressions, and how this application was not intrusive or in a total of 20 individuals.

Hence, the model for navigating a CD is based on two expressions:

- Open and close your eyes (move the seat forward or stop);
- Turning the head to the left or right (moves the seat through 90° about the axis thereof)

This work aims to present the architecture to detect the action of opening and closing the eyes based on HCC characteristics and evaluate the performance of the detector according to efficiency and effectiveness.

The effectiveness of the detector was tested in the prototype, evaluating their response to commands, since its effectiveness was evaluated from:

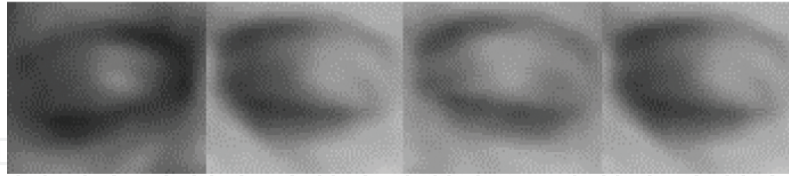
- Target image database construction considering regionalities and Brazilian cultural diversity;
- Manipulation of the various training variables;
- Handling and controlling the ideal minimum number of neighbors of the detector, reducing the number of false positives;
- Increased response or equal to 98% accuracy.

According to [11], the model [12] uses a classifier cascade and features of HCC. It works great on face detection and has already become standard for its high hit rate and low false positive rate.

The model [12] is adaptive and widely used by its degree of robustness and speed. In our implementation, the detection is based on HCC, with parameters evaluated, and a positive bank with about 10,000 images of frontal eyes closed, as shown in **Figure 2**.

This chapter is organized as follows: Section 2 presents the state of the art related to work developed here as well as our contribution. Section 3 describes the materials and the method

used for the realization and implementation of the architecture. Section 4 presents the tests and results. Section 5 in turn brings the conclusions and future work.



**Figure 2.** Sample det positive.

## 2. State of the art

According to [2], there are more than 35 CRI projects around the world, all of these projects vary in many ways. Initially, they include the project IntellWheels [2] which aims to create a development platform for intelligent wheelchairs, facilitating the design and testing of new methods and techniques for the CRIs, which takes as its premise aspects of low cost, comfort, and ergonomics.

Elizabeth and Roger [13] propose a seat which makes use of the head movement to perform navigation with a video system and a helmet with sensors. The study by [14] provides that the construction of an ellipsoidal 3D model of the head interprets the flow of this movement and establishes, according to the author, a more effective methodology than 2D approach.

Based on the Kalman filter to predict the possible movement of the head [15], attempts to reduce the action of the head movement assuming that with few moves the chair can now follow a route estimated by this position. This action, according to the author, reduces user effort.

Taylor and Nguyen [16] and Nguyen et al. [17] use similar ideas to use sensors placed on the user's head (especially accelerometers) to detect motion in a system designed to free platform. Wei-Kai and Isaac [18] address the recognition of facial gestures in interactive environments and aims to recognize the gestures through points of interest in the face of the individual by using Active Shape Models to extract face features to then assemble a 3D model in line with Actinos Facial Coding System (FACS).

Manogna et al. [19] create a system to directly control the engine speed based on the movement of the head. The device is fixed on the patient's head and sensor-based motion produces CRI. The same study was performed by [20], who evaluated a similar device in a closed environment. It describes the difficulty in understanding the face and his intentions and suggests using the same type of sensor.

The study by [19, 20] has been updated with the study by [21] that detects the opening and closing of the mouth for extra movement and studying the problem of head ergonomics. Zheng et al. [5] describe the Lucas-Kanade algorithm for detecting facial movement and discusses a method efficiency, accuracy, and response time.

Control of the CRA's status based on the opening and closing of the eyes is a simple method used to control the same as the state of the art analyzed but widely applied in research-related fatigue drivers to avoid traffic accidents.

Martin and colleagues [22, 23] describe a system that locates and tracks your eyes for algebraic operations on the face of the image extracted by the method of [20], which notably is also used in [24–30]. In all, eyes are found from the face detection which reduces the search space and facilitates significantly the detection task to then perform detection. Some utilize filters that improve the detection condition.

It is noticed that the utilization of computer vision for CRI movement was gradual. It was also almost constant that the use of accelerometers or cameras on the individual's head makes the equipment very intrusive and requires specific training to use the hardware and therefore difficult to adapt. The hit rates always fluctuate above 90%, but these mentioned, none is to analyze the influence of detector creation parameters. The majority does not have a specific database for the purpose and generally uses the same standard database available on the Web.

The contribution of this work is focused on the identification of the influence of the classifier training parameters and the relationship between complexity and efficiency/effectiveness of the detector.

Formation of a database of 10,000 images and analyze the influence of regionalism in the detector as opposed to embedded detectors in public libraries. The analysis of processing time both the detector's response to its creation.

### **3. Materials and method**

The model for the detection of eye state uses a simple idea. If a detector can detect a human eye on a front face then it must also recognize one eye closed. To this end, a trained classifier cascade is used to identify this object in the input image. The choice of [12] model mainly provides for its simplicity, speed of execution, and the outstanding performance [8]. The method basically combines four key concepts:

- Rectangular features called Haar features;
- Full image;
- Learning algorithm—AdaBoost;
- A classifier cascade.

The combination of these ideas permits simultaneous selection of key features and trains the classifier cascade, and then next steps will be described.

#### **3.1. Features of Haar**

The Haar features encode the existence of contrasts between the targeted regions of the image. A set of these resources is used to encode the contrasts displayed by a human face, on this

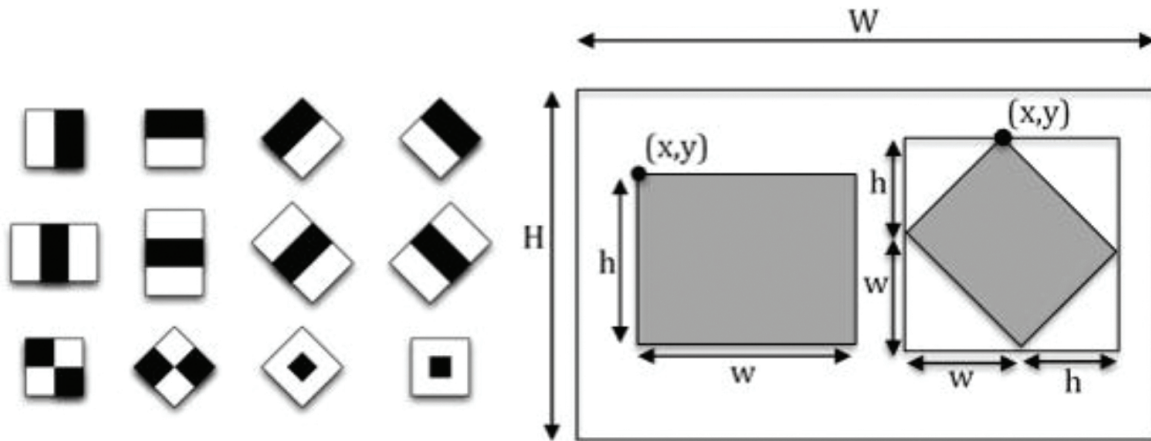


work, eyes closed, and their spatial relationships. These characteristics are called Haar, because its concept is similar to the coefficients of the Haar wavelet, set in a detection window  $W \times H$  pixels according to the formula:

$$\text{features} = \sum_{i=1}^N \omega_i \text{RecSum}(r_i) \quad (1)$$

where  $\omega_i$  is arbitrarily singling as weight factor, and  $\text{RecSum}(r_i)$  is the sum of the intensities of the pixels, which was described by [12] as a full image. The rectangle  $r_i$  and described as a function of five parameters:  $r = (x, y, w, h, \varphi)$ , where  $x, y$  are the coordinates of the top position of the pixel array,  $w$  and  $h$  define the dimensions of the rectangle, and  $\varphi = \{0^\circ, 45^\circ\}$  represents the degree of rotation.

The presence of a characteristic Haar, **Figure 3**, is determined by subtracting the average pixel value of the region by the average pixel value of the clear area. If the difference is above a threshold (set during learning), the characteristic is present.



**Figure 3.** Some features of Haar and the detection window.

Paul and Michael [12] reported the fact that the choice of the characteristics of use, rather than models based on pixel points of the image statistics, is important because of the benefits of the ad hoc domain knowledge, which can be extracted knowledge hidden in images hardly found in a finite set of training.

In the case of blink detection, this fact is used to represent the approximate information and also related to the test image backgrounds. This knowledge becomes very thin with respect to open and closed eye, the use of the two is hardly found in other models with appearance-based approach.

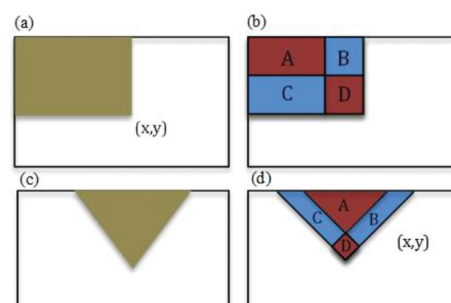
In general, therefore, the characteristics are nothing more than rating information for a set of light intensity of a pixel. This process consists of the sum of the intensity of the pixels of white

regions of characteristics and the intensity subtracted from the sum of the gray balance of the image. The results are used as the characteristic value of a given location and can be combined to form weak hypotheses in the images [31].

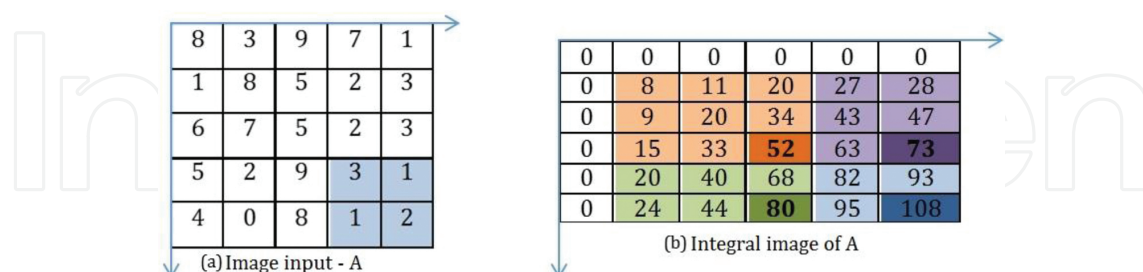
Typically, the model adopts the rectangles seen in **Figure 3** and determines the presence or absence of thousands of Haar features in each image position and with different scales, Paul and Michael [12] used a technique called integral image.

### 3.2. Full image

The complete image created from the original image, a new representation of the image, simply sums the values of each pixel to the left and above, inclusively. The idea to use this representation is to increase the speed of feature extraction, as any rectangle of an image can be calculated by means of this idea. Only four indices are required to calculate any rectangle, and as an immediate consequence, one needs only one pass for desired data in subregions of an image, see **Figures 4** and **5**.



**Figure 4.** Representation of full image: (a) area calculation, (b) sum of areas  $A-B-C+D$ , (c) rationed area calculation, (d) quick sum  $A-B-C+D$ .



**Figure 5.** Integral calculus image representation. It should be noted that the sum of the region (a) is equal to seven in (b) represented as  $108-73-80+52$ .

### 3.3. Adaboost

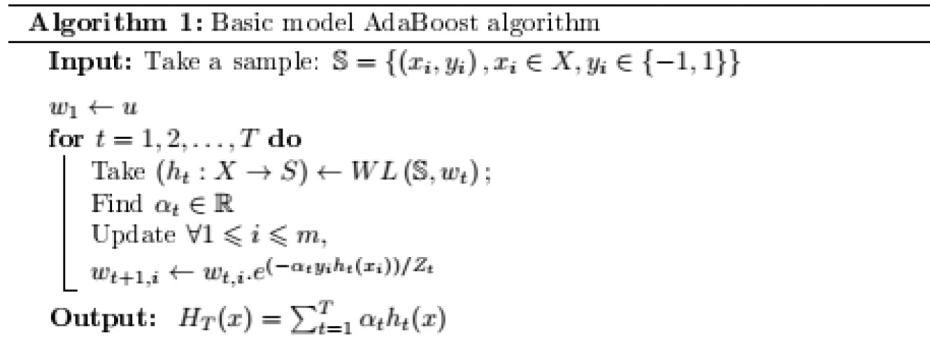
The source Boosting problem in computational area, known as machine learning, can be exposed informally as follows: suppose there is a sorting method which is slightly better than



a random choice for any distribution  $X$ , called weak learner or weak classifier. The existence of a classifier weak implies the existence of a strong classifier (strong learner), with small error on the entire space  $X$ .

In statistics, it is asked if given a reasonable estimation method, you can get a method close to great. This problem was solved by [32], which presented an algorithm that transforms a weak classifiers into a strong classifier.

From then on, several algorithms were developed within the context of boosting. One of the most recent and successful algorithms is known as AdaBoost that comes from the fact that boosting generates in every step a distribution on the observations of the sample and gives greater weight (most likely to be in the disturbed sample) to misclassified observations in the previous step. The basic algorithm is shown in **Figure 6**.



**Figure 6.** Adaboost algorithm.

In this sense, AdaBoost is focused on the bad ratings, or else the data difficult to classify, and this is the main feature of this algorithm: minimize error over a training set. One of the advantages of AdaBoost, as studied by [33, 34], is the existence of other parameters, in addition to  $T$  shifts, to improve learning.

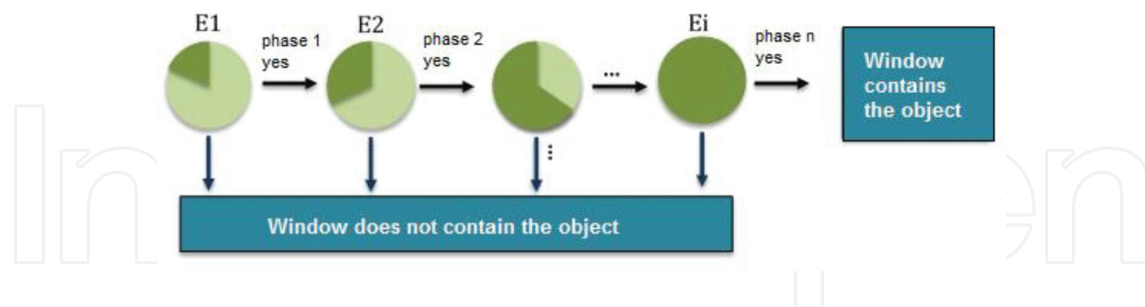
The result, after successive iterations of the algorithm, a set of hypotheses with weights wherein those having lower classification errors become more important, is called strong hypothesis or strong classifier.

### 3.4. Classifier cascaded

Increasing the speed of a classification task, in general, results in an increase in errors associated. However, for this purpose to be effective, we would have to reduce the number of evaluation of the weak classifiers, which would result in a loss in accuracy of the system. So Paul and Michael [12] propose a degenerative tree decision, decision stump, the structure that contains the binder thread from general to more specific, according to which the first cascade levels are not very accurate, although able to sort a large number of samples with a small amount of characteristics.

The use of cascade is characterized by the fact that, in an image during a detection task, the most sub-window analyzed by classifier is rejected. For this reason, a generalization in the

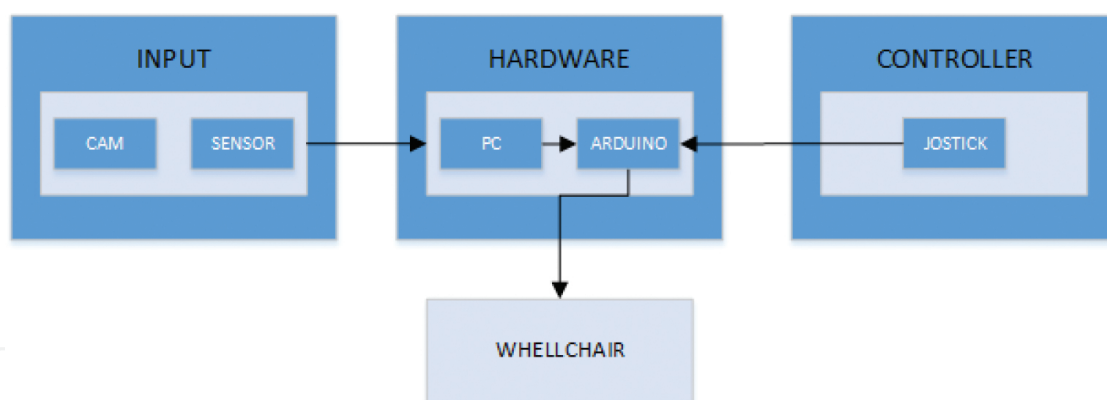
early stages should be high enough to avoid the transition to subsequent stages of sub-windows classified as false positive [35], as shown in **Figure 7**.



**Figure 7.** Estructure of the classifier.

### 3.5. The chair

One of the general proposed projects involves low cost as the minimum use of equipment is attached to the chair. This comprises only a webcam coupled to the PC and a Kinect® sensor (see **Figure 8**) coupled to the associated chair or independently allows the model to obtain various environmental information capable of determining actions of the chair, for example the sudden passage of someone in front of the same or the proximity of an obstacle.



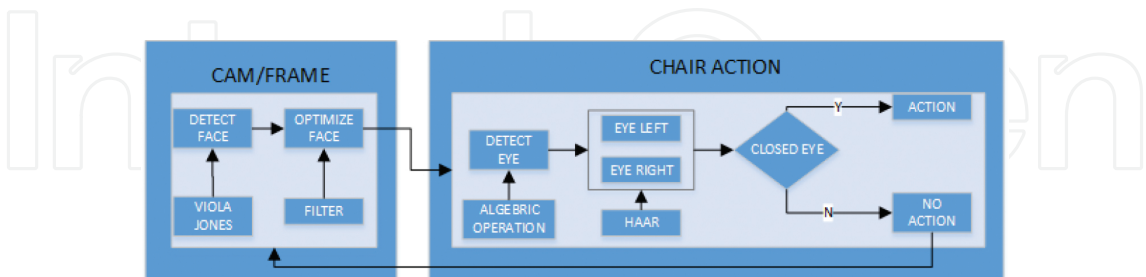
**Figure 8.** Chair and its embedded hardware.

A notebook with processor 1.6 GHz dual-core Intel Core i5 (Turbo Boost up to 2.7 GHz) with 3 MB and 4 GB of memory was used. The connection between the PC and the chair uses an ATmega328 microcontroller board (arduino UNO), which has 14 digital pins input/output, 6 analog inputs, and USB connection.

### 3.6. Architecture

The software implements the model, which seeks, from the detection of the eye status (open, closed) to inform the chair the action that should be taken especially to move forward or stop. The command “move on” triggered when the user closes his eyes unnaturally or stays with

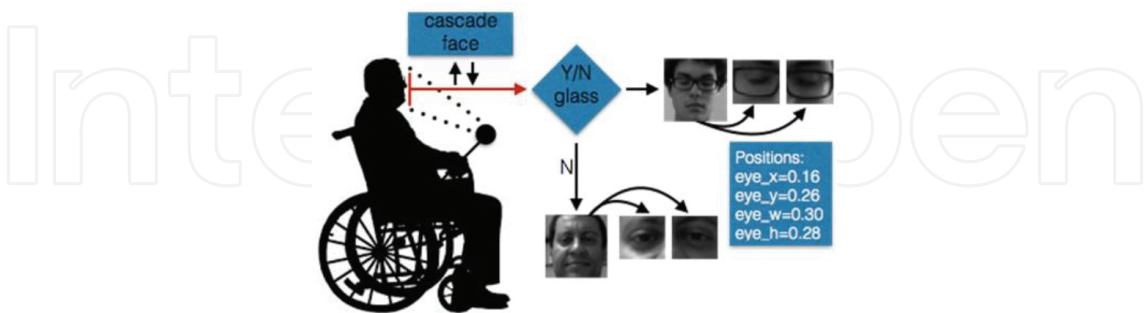
the eye closed for a period of time greater than 2 s, in about 10 frames, is analyzed, and if the detection occurs in a percentage above 90%, a command is sent, otherwise nothing is done. The command acts in a similar way as opposed to the motion to move forward as a way to clear up this script diagrammed in **Figure 9**. It was found that according to [36], the human eye takes 280 ms to blink.



**Figure 9.** How to detect closed eyes.

In the proposed architecture, **Figure 9**, CAM/frame image captures through a common webcam, performs the face detection, and focuses on this area of interest; then the image is treated to minimize or neutralize lighting noise and send the information for treatment. In CHAIR ACTION, capturing is carried out prior to information and detects by means of an algebraic operation described in [36], the region where the eyes in the face are properly separated and sent to analyze the state close left eye (CL) and close right eye (CR). In the first step, the image is captured and represented in gray tones; there are 28 frames per second. In each, we use facial detector defined in [12] widely used and recognized efficiency, according to [37]. Around this region of interest are made two image optimization operations:

- Inversion;
- Retinal filter.



**Figure 10.** Eye detection structure.

Filtering through the algorithm in [38] called retinal filter cancels much of the image distortion and improves cleaning detail, even in low light (<100 lux) also keeps the naturalness. After correction of the image, an algebraic operation is made to detect the eyes of the person in the image. The completion of this step proved to be quite efficient with the use in [27] or [28] due

to the fact that even after correction of the image both methods provide great instability regarding the detection of the eyes, even using different parameters as those used for the same. The extraction assumes that the person is in front of the camera and below the horizontal line of the eyes, while maintaining a fixed distance to the camera in a variable angle between  $30 < \theta < 70$  as described in **Figure 10**.

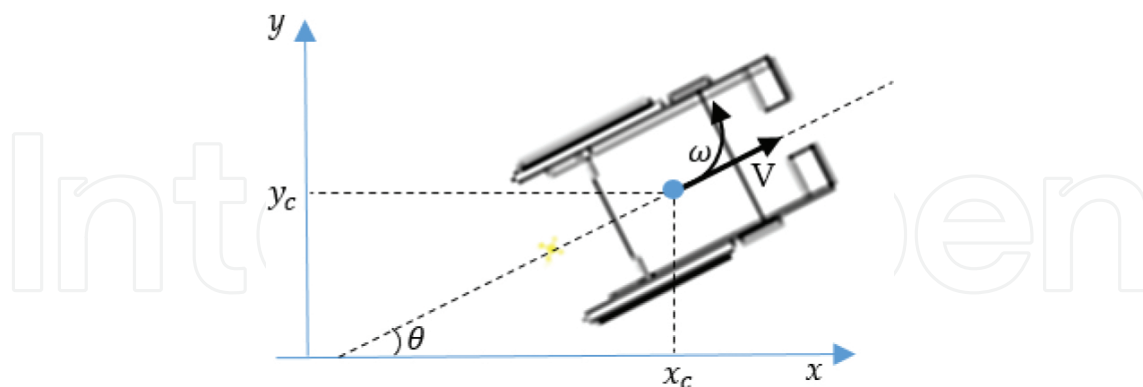
### 3.7. Kinematic model and motion control

The kinematic model of the chair is presented as a process in which each wheel contributes both to the movement of the chair as to his/her mistake notably associated with obstacles and soil deformation. While our research ambience is indoors and appropriate to the patient, these errors can be minimized with the use of inexpensive sensors that can identify and predict possible problems in navigation.

Here we describe the mathematical idea of the chair movement that follows a traditional model of representation of the world. This model, whose movement and orientation are performed by two independent actuators, considers the rectangular object moving at speed  $V$ . The state plan of the chair in the Cartesian plane  $(x, y)$  is defined by the vector:

$$(x_c, y_c, \theta, v_c, \omega_c)^t \quad (2)$$

where  $x_c$  and  $y_c$  are the coordinates of the central point of the wheel axle,  $\theta$  is the angle formed between the base of the chair  $C(x_c, y_c)$ ,  $v_c$  is the linear velocity at the point  $C$ , and  $\omega_c$  is the angular velocity,, as described in **Figure 11**.



**Figure 11.** Cartesian representation of the chair.

The chair should move only in the direction normal to the axis of the driving wheels and can restrict the analytic relationship:

$$y' \cos \theta - x' \sin \theta \quad (3)$$

Based on the information that is obtained during the navigation, the calculation of the linear speed of each wheel is deducted for subsequent adjustment and especially in cases of unevenness through a relation between the number of pulses of the encoder  $N$  and its sampling period  $T$ .

$$v = \frac{N\pi D}{R_e T} \quad (4)$$

where  $v$  and  $D$  are, respectively, the linear velocity and the diameter of the wheel, and  $R_e$  is the encoder resolution. A possibility of representation of state variables is based on the speed at the contact point between the right wheel ( $v_D$ ) and left ( $v_E$ ) with the floor.

$$(x_c, y_c, \theta, v_D, v_E)^T \quad (5)$$

The choice of this form of representation is essentially the ease of mediating these quantities by odometry system.

Considering the continuous system, we have:

$$\begin{cases} v_D = \left(L + \frac{b}{2}\right) \omega = r_D \omega_D \\ v_E = \left(L - \frac{b}{2}\right) \omega = r_E \omega_E \end{cases} \quad (6)$$

Besides that

$$\begin{cases} v_D = v + \left(\frac{b}{2}\right) \omega \\ v_E = v - \left(\frac{b}{2}\right) \omega \end{cases} \quad (7)$$

Also,

$$\begin{cases} v_D + v_E = 2\omega r = 2v \\ v_D - v_E = \omega L \end{cases} \quad (8)$$

or alternatively

$$\begin{cases} v = \omega_D \frac{r_D}{2} + \omega_E \frac{r_E}{2} \\ \omega = \omega_D \frac{r_D}{2} + \omega_E \frac{r_E}{b} \end{cases} \quad (9)$$

Consider  $b$  the distance between the contact points of the wheels,  $r_D$  radius of the wheels,  $L$  the distance from point  $C$  to the center of rotation of the chair, and  $\omega_D, \omega_E, \omega$  the angular velocities of the right, left, and center wheel movement of the chair.

From the speeds of the wheels of the robot, we may calculate the linear and angular velocities.

$$\begin{cases} v = \frac{v_D + v_E}{2} \\ \omega = \frac{v_D - v_E}{b} \end{cases} \quad (10)$$

To find out its position in the reference plane, we must know the chair state space, and how it will evolve over time with the  $v_D$  and  $v_E$  speeds.

Considering the condition called no slip can describe the kinematic equations of motion of point  $C$  with respect to the linear ( $v$ ) and angular ( $\omega$ ) velocity [4, 26]:

$$\begin{cases} x_c = v \cos \theta \\ y_c = v \sin \theta \\ \theta = \omega \end{cases} \quad (11)$$

or matrix form:

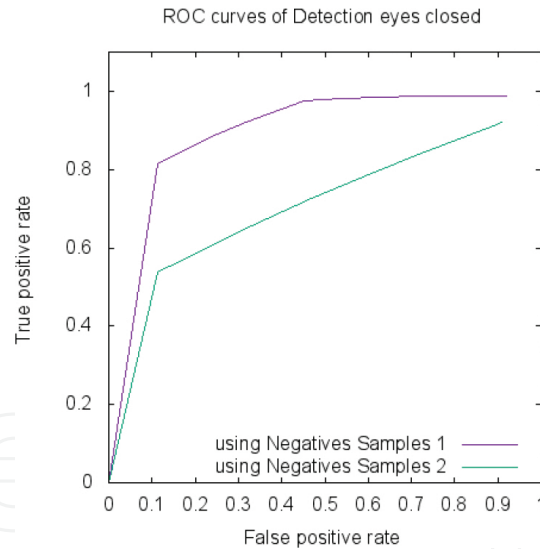


$$\begin{bmatrix} x_c \\ y_c \\ \theta \end{bmatrix} = \begin{bmatrix} \cos \theta & 0 \\ \sin \theta & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v \\ \omega \end{bmatrix} \quad (12)$$

#### 4. Testing and results

The performance of the closed eye detection system was built based on a set of 10,000 images of various persons. Images were acquired in a controlled lighting environment with  $2048 \times 1536$  pixels in JPG format. For accurate detection classifier, various parameters can be changed during the training process. The influence of these parameters changes the complexity of weak classifiers, and therefore aspects as positive and false positives are influenced.

The standard size entry was set to  $24 \times 24$  pixels. All images in our base were taken in the same, uniform background. Immediately, we wanted to see if the negative assembly constructed from the same occluded images with faces is sufficient to distinguish between open and closed eyes; using the classifier with default parameters, the following results were obtained, as described in **Figure 11**.



**Figure 12.** Comparison between detector 1 taken with random sample negatives and diverse background and using samples only performing the occlusion of the region of interest.

This analysis would serve to decrease the cost of both image search optimization and scheduling but has not proven satisfactory for such small images. Then the other negative training set was created, gathered randomly about 9000 different images that do not contain any references to human eyes. **Figure 12** shows some examples of negative training sets. As Rainer and Jochen [39] showed, the version of AdaBoost gave best results for facial detection, Gentle AdaBoost, with a ratio of false positives required cascades set to  $10e^{-6}$ .

Due to small size of the training image, we limited rectangle Haar that was used. Although the number of ways in which the rectangles may be arranged is large, for practical reasons, we limited the time with the following steps:

1. Only Haar-like with two, three, and four rectangles Was considered;
2. The size of the model of Haar features was set at a maximum of  $5 \times 5$  and  $3 \times 3$  pixels at least;
3. All rectangles that contribute to the unique Haar features were of the same size.

A total of 408,564 characteristic Haar were obtained by imaging under the above conditions with a satisfactory number of resources. For this problem, four detectors were used which were constructed and differed primarily by the type of boost used and parameter variation as described in **Table 2**, wherein MinHitRate is the minimum hit rate desired for each phase of the classifier; MaxFalseAlarm is the maximum false alarm rate desired for each phase of the classifier; Nstages is the number of cascaded stages; BTYPE is the kind of boost used (type of boosted classifiers: DAB— Discrete AdaBoost, RAB—Real AdaBoost, LB—LogitBoost, GAB— Gentle AdaBoost); WTRate is the cutting line and the weight used in the boost; and Wcount is the maximum count of false trees for all stages of the cascade.

Parameter	Sorter 1	Sorter 2	Sorter 3	Sorter 4
MinHitRate	0.9–0.999	0.9–0.999	0.9–0.999	0.9–0.999
MaxFalseAlarm	0.1–0.5	0.1–0.5	0.1–0.5	0.1–0.5
Nstages	20	25	20	25
Btype	GAB	RAB	LB	DAB
WTRate	0.95–0.98	0.95–0.98	0.95–0.98	0.95–0.98
Wcount	100	100	100	100

**Table 2.** Presentation of parameters and classifiers.

As shown in **Figure 13**, the results are established with great parameter, and the results are described in **Table 3**.



**Figure 13.** Negative sample.

Parameter	Sorter 1	Sorter 2	Sorter 3	Sorter 4
MinHitRate	0.999	0.987	0.985	0.999
MaxFalseAlarm	0.5	0.5	0.4	0.5
Nstages	20	25	20	25
Btype	GAB	RAB	LB	DAB
WTRate	0.97	0.98	0.95	0.95
Wcount	100	100	100	100
Processing time	4 days	5 days	4 days	7 days

Table 3. Presentation of results and better processing time.

Figure 14 shows the difference between the result of the public detector and the detector with better response. Both have a good recognition rate, but positive for higher rates against false positives makes the classifier "olhosfechadosGAB" provide better performance.

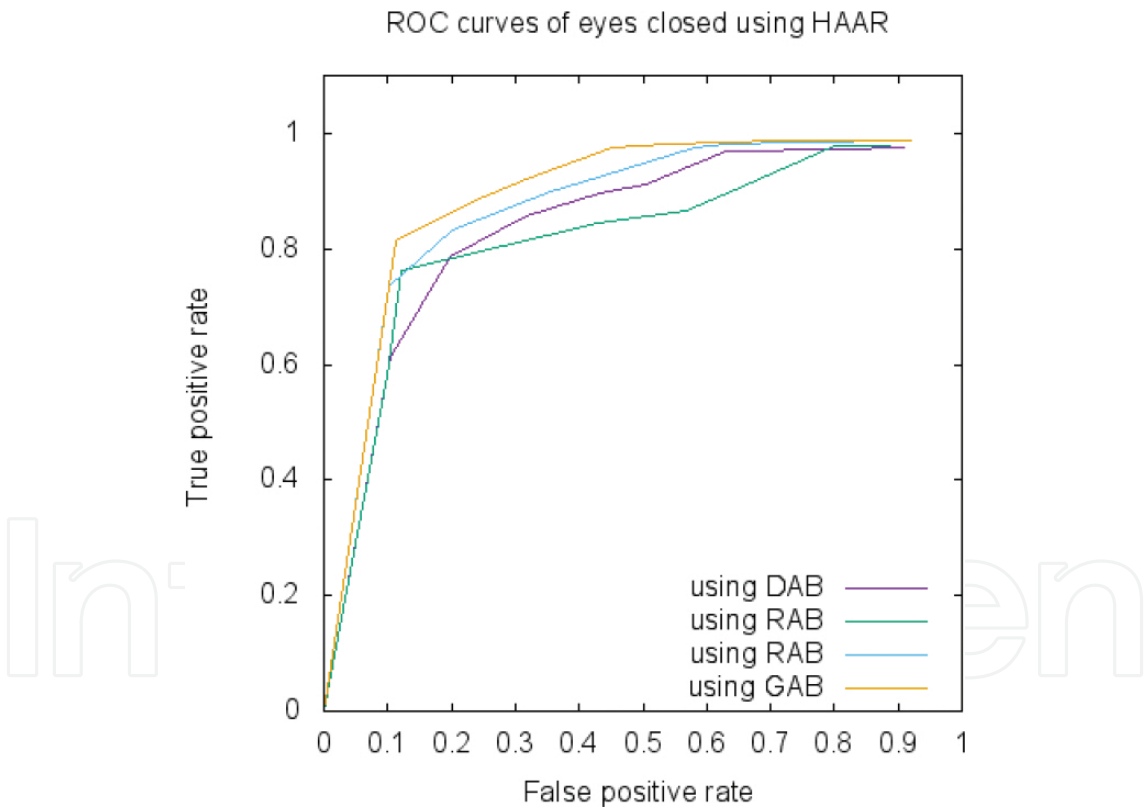
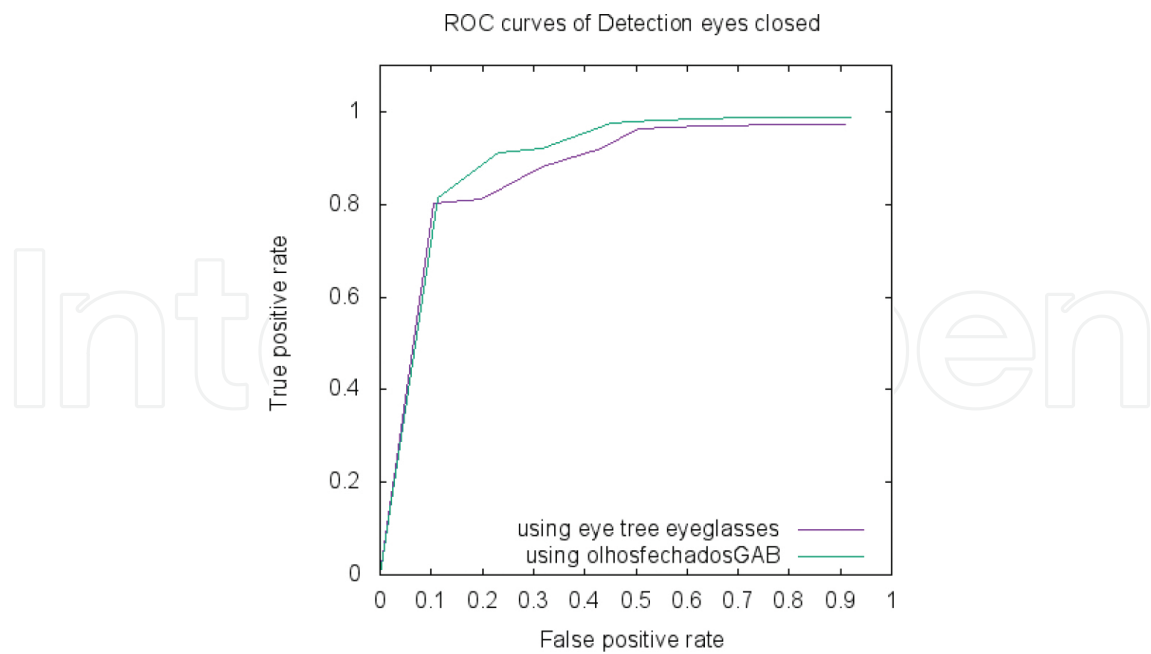


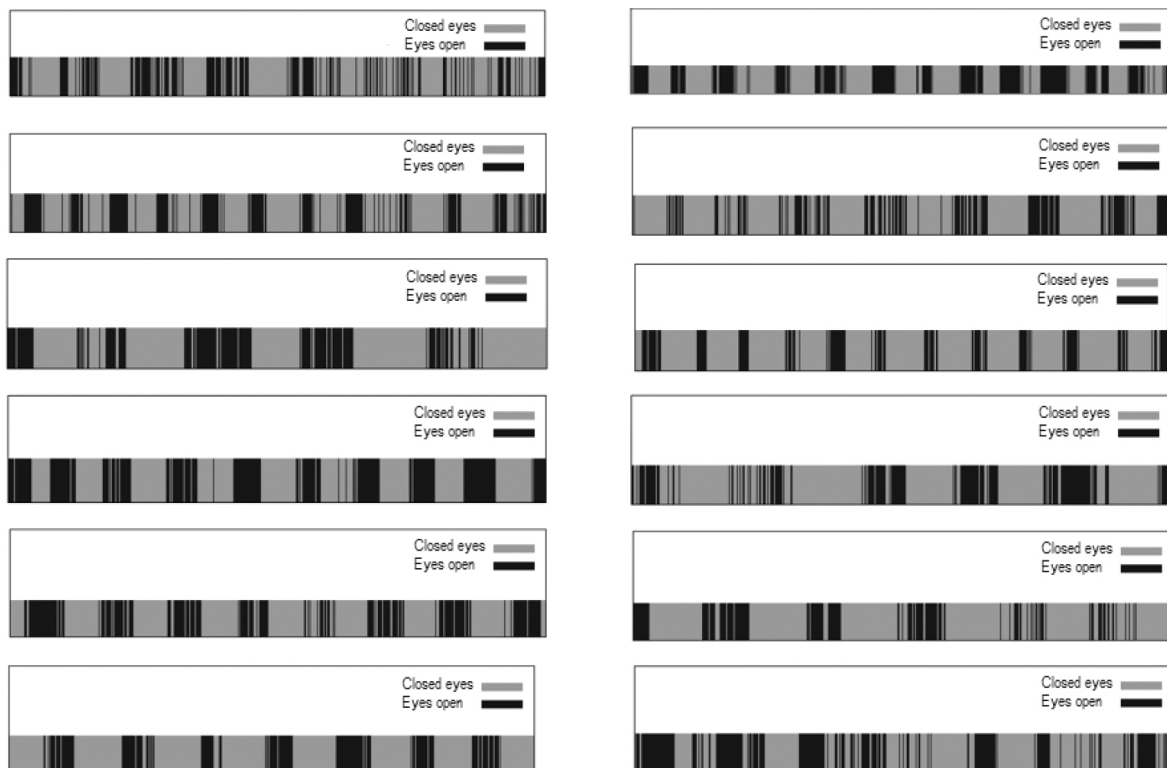
Figure 14. ROC graph with the best parameters and analysis for different versions of AdaBoost.

Simulating what will happen directly in the chair is crucial. Therefore there were two tests: the first with 12 volunteers, conducted during the same period of time (2 min) recording a video with one's image of the front face held at specified times to opening and closing the eyes, with the following results described in Figure 15.



**Figure 15.** Comparison of our classifier and public distribution.

In the second test, the same group of volunteers held navigation CD from point A to point B, which would use a region of 1 square meter stopping place, as shown in **Figure 16**.



**Figure 16.** Result detection of volunteers.

Volunteers can perform spins of the head that do not exceed 30 degrees, in an environment with good lighting conditions (above 200 lux). In carrying out the detection was operated with frame size 640 by 320 with 28 FPS. In all, the volunteers were able to make the journey without difficulty with stopped or early movement. During the test, we observed the total time hardware response to commands, and these times ranged from 250 to 300 ms.

## 5. Conclusion

The tests clearly demonstrated that HCC can be successfully used in a blink detection system, and the combination of the classifier set to closed eyes and face resulted in a fast and efficient system.

The trained detector (one classifier) with the parameters described in **Table 2** exceeded the detector OpenCV framework proposed by both the rate of detection and computational efficiency. In this study, we were able to detect approximately 98% of the results with about 9% false positives.

The results showed that the use of regionalized data enables more efficient detection. We observed that the detector does not fail to be submitted to the people with whom he/she was trained. To set a robust system, the patient must have a face image trained classifier with a significant number of possessions.

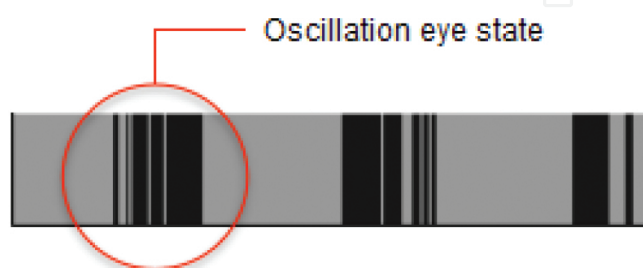


**Figure 17.** Conducted in chair.

The average error was 0.058, while applying the minimum neighboring detector returned a number of 90% positive, while for the maximum number of neighbors obtained 98% as detection windows scan across the image with the maximum neighboring the intersection between the windows is unique and therefore there is no possibility of true or false.

Another crucial point is the detector oscillation while the individual closes and opens his eyes. This set of errors stabilizes in a few milliseconds (would not be a problem if this time (**Figure 17**) did not result in "leaps" in the chair). The solution to this problem was given by using a waiting time for stabilization of detection (about 1 s), and thereafter the chair had its start and stop smoothly performed.

Using our PC, it was possible to obtain an average rate of 280 ms detector response, and the architecture proved to be quite stable, no crash or time-consuming to user responses. Individuals engaged in the work reported that it would not require extensive training to use the chair and (after 30 min of test) and did not experience discomfort when using the system (**Figure 18**).



**Figure 18.** Detector response analysis.

As future work, we highlight the implementation of the navigation control with the head that will allow the user to perform turns and spins with the chair and the analysis of all these ideas applied to a larger architecture and its integration with other low-cost sensors that allow to bypass obstacles.

## Author details

Marcos Figueredo<sup>1\*</sup>, Alexandre Nascimento<sup>2</sup>, Roberto L.S. Monteiro<sup>1,3</sup> and Marcelo A. Moret<sup>3,4</sup>

\*Address all correspondence to: [mbfigueredo@uneb.br](mailto:mbfigueredo@uneb.br)

1 Department of Exact Sciences and Earth—DECET II, College of Information Systems—Alagoinhas, State University of Bahia—UNEB, Bahia, Brazil

2 College ÁREA1, Salvador, Bahia, Brazil

3 Doctoral Program in Computational Modeling and Industrial Technology, College of SEN-AI-CIMATEC Technology, Salvador, Bahia, Brazil

4 Department of Exact Sciences and Earth—DCET, Campus I, State University of Feira de Santana—UESF, Salvador, Bahia, Brazil



## References

- [1] Instituto Brasileiro de Geografia e Estatística—IBGE. Características da população e dos domicílios resultados do universo, Censo Demográfico 2010, vol. 1, p. 161, 2011.
- [2] BRAGA, Rodrigo António Marques et al. Plataforma de desenvolvimento de cadeiras de rodas inteligentes. PhD, Thesis, Faculty of Engineering, University of Porto, PhD Program in Computer Science, Porto, Portugal, set 2012.
- [3] HALAWANI, Alaa et al. Active vision for controlling an electric wheelchair. *Intelligent Service Robotics*, v. 5, n. 2, pp. 89–98, 2012.
- [4] SONG, You; LUO, Yunfeng; LIN, Jun. Detection of movements of head and mouth to provide computer access for disabled. In: 2011 International Conference on Technologies and Applications of Artificial Intelligence (TAAI). IEEE, 2011, pp. 223–226.
- [5] ZHAO, Zheng; WANG, Yuchuan; FU, Shengbo. Head movement recognition based on Lucas-Kanade algorithm. In: 2012 International Conference on Computer Science & Service System (CSSS). IEEE, 2012, pp. 2303–2306.
- [6] RUMÃO DE MELO, Valdenice. Avaliação da qualidade de vida de pacientes com lesão medular acompanhados em regime ambulatorial. Master's thesis - Federal University of Pernambuco. CCS. Neuropsychiatry and Behavioral Sciences, 2009.
- [7] CEZAR DA CRUZ, Daniel Marinho; AUGUSTO IOSHIMOTO, Maria Teresa. Tecnologia assistiva para as atividades de vida diária na tetraplegia completa c6 pós-lesão medular. *Magazine Triângulo*, v. 3, n. 2, 2011.
- [8] Learning OPENCV. Computer vision with the OpenCV library. Gary Bradski & Adrian Kaebler. O'Reilly, 2008.
- [9] BRAGA, Rodrigo António Marques et al. Concept and design of the intellwheels platform for developing intelligent wheelchairs. In: *Informatics in Control, Automation and Robotics*, Springer-Verlag, Berlin Heidelberg, 2009, pp. 191–203.
- [10] JIAN-ZHENG, Liu; ZHENG, Zhao. Head movement recognition based on lk algorithm and gentleboost. In: 2011 7th International Conference on Networked Computing and Advanced Information Management (NCM). IEEE, 2011, pp. 232–236.
- [11] JOLLIFFE, Ian. *Principal Component Analysis*. John Wiley & Sons, Ltd, 2002.
- [12] VIOLA, Paul; JONES, Michael. Rapid object detection using a boosted cascade of simple features. In: *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2001. CVPR 2001. IEEE, 2001, 1, pp. I-511–I-518.
- [13] DYMOND, Elizabeth; POTTER, Roger. Head movements for control of assistive technology. *Engineering in Medicine and Biology Society*. In: 14th Annual International Conference of the IEEE, 1992, 4, pp. 1527–1528.

- [14] BASU, Sumit; ESSA, Irfan; PENTLAND, Alex. Motion regularization for model-based head tracking. In: Proceedings of the IEEE International Conference on Pattern Recognition (ICPR '96). Vienna, Austria, 1996, 3, pp. 611–616.
- [15] KIRULUTA, Andrew; EIZENMAN, Moshe; PASUPATHY, Subbarayan. Predictive head movement tracking using a Kalman filter. IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics, v. 27, n. 2, pp. 326–331, 1997.
- [16] TAYLOR, P. B.; NGUYEN, H. T. Performance of a head-movement interface for wheelchair control. In: Proceedings of the 25<sup>th</sup> Annual International Conference of the IEEE EMBS, 2003, pp. 17–21.
- [17] NGUYEN, H. T.; KING, L. M.; KNIGHT, G. Real-time head movement system and embedded linux implementation for the control of power wheelchairs. In: 26<sup>th</sup> Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2004. IEMBS'04. IEEE, 2004, pp. 4892–4895.
- [18] LIAO, Wei-Kai; COHEN, Isaac. Classifying facial gestures in presence of head motion computer vision and pattern recognition – workshops. IEEE Computer Society Conference on CVPR Workshops. 2005, p. 77.
- [19] MANOGNA, S.; VAISHNAVI, Sree; GEETHANJALI, B. Head movement based assist system for physically challenged. In: 4<sup>th</sup> International Conference on Bioinformatics and Biomedical Engineering (iCBBE), 2010, pp. 1–4.
- [20] WEI, L.; HU, H.; LU, T. & YUAN, K. Evaluating the performance of a face movement based wheelchair control interface in an indoor environment. In: Proceedings of the IEEE International Conference on Robotics and Biomimetics, 2010, pp. 387–392.
- [21] LEE, Chan-Su; SAMARAS, Dimitris. Analysis and synthesis of facial expressions using decomposable nonlinear generative models. In: 2011 IEEE International Conference on Automatic Face & Gesture Recognition and Workshops (FG 2011). IEEE, 2011, pp. 847–852.
- [22] ERIKSSON, Martin; PAPANIKOTOPOULOS, Nikolaos P. Eye-tracking for detection of driver fatigue. In: IEEE Conference on Intelligent Transportation System, 1997. ITSC'97. IEEE, 1997, pp. 314–319.
- [23] ERIKSSON, Martin; PAPANIKOLOPOULOS, Nikolaos P. Driver fatigue: a vision-based approach to automatic diagnosis. Transportation Research Part C: Emerging Technologies, v. 9, n. 6, pp. 399–413, 2001.
- [24] LIN, Chern-Sheng; CHANG, Kai-Chieh; JAIN, Young-Jou. A new data processing and calibration method for an eye-tracking device pronunciation system. Optics & Laser Technology, v. 34, n. 5, pp. 405–413, 2002.
- [25] JI, Qiang; YANG, Xiaojie. Real-time eye, gaze, and face pose tracking for monitoring driver vigilance. Real-Time Imaging, v. 8, n. 5, pp. 357–377, 2002.

- [26] ARAI, Kohei; MARDIYANTO, Ronny. Eyes based electric wheel chair control system. *International Journal of Advanced Computer Science and Applications*, v. 2, n. 12, 2011.
- [27] CHAU, Michael; BETKE, Margrit. Real time eye tracking and blink detection with USB cameras. *Boston University Computer Science*, v. 2215, n. 2005–2012, pp. 1–10, 2005.
- [28] FAZLI, Saeid; ESFEHANI, Parisa. Tracking eye state for fatigue detection. In: *International Conference on Advances in Computer and Electrical Engineering (ICACEE 2012)*. 2012. pp. 17–20.
- [29] ALSHAQAQI, Belal et al. Driver drowsiness detection system. In: *2013 8th International Workshop on Systems, Signal Processing and their Applications (WoSSPA)*. IEEE, 2013, pp. 151–155.
- [30] SAINI, Vandna; SAINI, Rekha. Driver drowsiness detection system and techniques: a review. *International Journal of Computer Science and Information Technologies*, v. 5, n. 3, pp. 4245–4249, 2014.
- [31] HJELMÅS, Erik; LOW, Boon Kee. Face detection: a survey. *Computer Vision and Image Understanding*, v. 83, n. 3, pp. 236–274, 2001.
- [32] SCHAPIRE, Robert E. The strength of weak learnability. *Machine Learning*, v. 5, n. 2, pp. 197–227, 1990.
- [33] NOCK, Richard; NIELSEN, Frank. A real generalization of discrete adaboost. *Frontiers in Artificial Intelligence and Applications*, v. 141, p. 509, 2006.
- [34] GAO, Wei; ZHOU, Zhi-Hua. Approximation stability and boosting. *Algorithmic Learning Theory*, v. 21, pp. 59–73, 2010.
- [35] HORTON, Michael; CAMERON-JONES, Mike; WILLIAMS, Raymond. Multiple classifier object detection with confidence measures. In: *AI 2007: Advances in Artificial Intelligence*. Springer, Berlin, Heidelberg, 2007, pp. 559–568.
- [36] EKMAN, Paul; ROSENBERG, Erika L. What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS). Oxford University Press, USA, 1997.
- [37] TURK, Matthew; PENTLAND, Alex. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, v. 3, n. 1, pp. 71–86, 1991.
- [38] BENOIT, Alexandre et al. Using human visual system modeling for bio-inspired low level image processing. *Computer Vision and Image Understanding*, v. 114, n. 7, pp. 758–773, 2010.
- [39] LIENHART, Rainer; MAYDT, Jochen. An extended set of haar-like features for rapid object detection. In: *Proceedings of the 2002 International Conference on Image Processing*. IEEE, 2002. pp. I-900–I-903, vol. 1.