

Real-Time Gaze-Tracking Embedded-System

Cristian E. Ordoñez^{*†}, Eduardo L. Blotta^{*} Juan I. Pastore^{*}

^{*}Laboratorio de Procesamiento Digital de Imágenes, Facultad de Ingeniería,
ICYTE-CONICET, U.N.M.D.P., Mar del Plata.

[†] Departamento de Informática, Facultad de Ingeniería, U.N.M.D.P.
ordonezc@fi.mdp.edu.ar

Abstract—This paper proposes a video-based Real-Time Gaze-Tracking system which allows to identify gaze commands to drive an electric powered wheelchair. The main aspect of the method is its ability to be implemented in a portable computing system, reduced both in processing capacity and in RAM memory. The Gaze-Tracking system is based on the Hough Transform, the use of shape features and the dark pupil effect. The method performance was validated using an own database, built under different ambient lighting conditions.

Index Terms—Real-Time-Processing, Embedded System, Hough Transform, Gaze-Tracking, Shape features.

I. INTRODUCTION

Severe injuries in the upper portion of the spinal cord, age, or diseases that affect motor neurons (MNDs), may prevent a person to perform tasks such as driving a common electric-powered wheelchair (EPW), usually controlled by a hand joystick.

Nowadays, there is a wide field of human-computer interfaces (HMIs) that allow to control an EPW without the need of using hands. Within this type of technology there are systems based on Electromyography (EMG) signals, Electroencephalogram (EEG) signals, sip and puff (SNP) sensors, tongue movement sensors, chin movement joysticks, and eye tracking methods.

EMG signal-based systems employ a sensor that converts muscle activity, usually from some muscle of the user's face, into small electrical impulses that are used to control an EPW [1] [2].

Brian-Computer interfaces (BCI) based on EEG signals have as disadvantage the high cost that involves linking the human brain to the device.

In SNP systems the user gives commands to the wheelchair by “sipping” and by “puffing” on a pneumatic tube like the one used in [4]. This method responds to the amount of pressure applied to the pneumatic tube and whether the sign of the pressure is negative or positive. These systems require a certain precision in each sip and puff and are not suitable for people with reduced respiratory capacity.

Interfaces that allow to control an EPW through the tongue are usually invasive, since they use an intra-oral device that registers the variations of inductance due to the movements of a ferromagnetic piercing that is placed in the user tongue [6] [7].

Chin-based devices use a force-sensing joystick usually controlled by neck flexion, extension and rotation [8].

Finally, eye tracking-based systems include a wide variety of methods. Within this type of systems are the devices based on sensors, which employ electrodes placed around the eye (Electro-oculogram, EOG), and video-based systems, where a video camera captures eye movements and a computer saves and analyzes the information sent. An advantage of EOG is its low influence on lighting conditions, however, it is an invasive technique that may be considered impractical for everyday use [9].

Many researchers present video-based eye tracking systems, where they implement techniques based on pattern recognition, corneal reflection points, dark-bright pupil effect through infrared illumination, eye models and hybrid techniques [10]. The implementation of these techniques, when computing power resources are available, does not represent challenges. However, when the design is oriented to a portable battery-powered computing system, the challenge is greater as it seeks to maximize computing efficiency in order to minimize consumption.

This work is an extension of [11]. An hybrid method based on the Hough Transform (HT), the use of shape features and the “dark pupil” effect is proposed to implement a wheelchair control embedded system using a portable low-cost, low-power computing system, based on a 32-bit microcontroller. In addition to the advances made in the pupil detection method, an infrared light source is added to the system in order to improve the pupil-iris contrast and a geometric criterion is established to identify movement commands for the EPW.

The method proposes adjustments that improve the HT performance in the pupil detection and also reduce the number of operations to be executed.

This paper is organized as follows: section II presents the theoretical bases, eye safety issues using infrared light and hardware description, section III details the method and section IV shows the results.

II. MATERIALS AND METHODS

A. Hough Transform

When extracting shape information of an object in an image, the most relevant information is often obtained around the boundaries or edges of the object [14], therefore the edges are used as a discriminative property. The Hough Transform (HT) is a global processing technique that allows to identify and locate objects whose shape is parameterized by the equation $g(v, c) = 0$, where v is a vector of coordinates and c is a vector

of coefficients. This technique is robust to noise presence and also to incomplete edges. HT turns a segmentation problem in the image space, in a peak detection problem in the parameters space. Paul Hough developed the HT in 1962 to detect straight lines [15]. Subsequently, Duda and Hart improved the Hough technique extending it to detect other types of parametrical curves [16]. Merlin and Faber [17] probed that the HT can be generalized to detect arbitrary shapes with a certain scale and orientation. Ballard eventually extended the scope of the work, formulating the Generalized Hough Transform (GHT) to efficiently detect arbitrary shapes with any orientation and scale [14] [18].

Starting from the premise that pupil has a circular shape, in this paper we propose to apply the Circle Hough Transform (CHT) in order to detect the pupil position. Although the premise of pupil circularity is not always satisfied, the CHT algorithm is good enough for our purpose. CHT is similar to the standard or linear HT, only the space parameters are different, corresponding in this case to the analytical model of the circle [19]:

$$(x - c_1)^2 + (y - c_2)^2 = c_3^2 \quad (1)$$

where (c_1, c_2) are the coordinates of the center of a circle of radius c_3 passing through (x, y) , on a cartesian coordinate system.

It is usual to use the parametric representation of the circle to implement the CHT [19]:

$$\begin{aligned} x &= c_1 + c_3 * \cos(\theta) \\ y &= c_2 + c_3 * \sin(\theta) \end{aligned} \quad (2)$$

where $\theta \in (0, 2\pi]$.

B. Shape features

The shape descriptor of a region is a measure of the degree of similarity with a given shape. The proposed method uses this measure as a criterion of discrimination between Hough candidates that produce local maxima in the accumulator space.

Most traditional shape descriptors are mere combinations of size parameters (area, perimeter and diameter) arranged in such a way that dimensions get canceled. Table I lists the most used shape descriptors. Each of these parameters captures an aspect of shape [12].

Shape features	Reference
$F = \frac{4\pi A}{P^2}$	Shape factor
$R = \frac{4A}{\pi D_{max}^2}$	Roundness
$R_a = \frac{D_{max}}{D_{min}}$	Aspect relationship
$E = \frac{l_f}{a_f}$	Elongation
$C = \frac{P_c}{P}$	Convexity

TABLE I: Shape features

Where:

A = area enclosed by the contour;

D_{min}/D_{max} = minimum/maximum diagonal of the area enclosed by the contour;

l_f = length: defined as the length along the axis of the curvilinear shape;

a_f = maximum width of the curvilinear shape;

P = perimeter;

P_c = convex perimeter: defined as the length of a convex line circumscribing the object to be measured.

C. Hardware Description

The proposed system runs on a *Discovery* kit, which is based on a *STM32F407VGT6* Micro Controller Unit (MCU), delivering 180 DMIPS @ 144 MHz, see Fig. 1. A base board, connected to the STM32F4 Discovery provides a micro SD Card slot and extension connectors for the LCD and camera boards. A digital camera board, featuring a 1.3 *Megapixel* CMOS OV9655 sensor and a 3"5 LCD board, with touch screen capability compose the whole system. More details on [24].



Fig. 1: Discovery Kit: main board, base board, camera and display



Fig. 2: Test bench

In order to evaluate the system performance a test bench was developed, as shown in Fig. 2. It allows to place the Discovery board and the LCD display to visualize the detection performed. In addition, it is possible to rotate and adjust the height at which the screen is located to improve the user comfort.

Regarding the Gaze Tracking glasses, an infrared light source was positioned next to the camera to illuminate the user's retina, as shown in Fig. 4.

Every camera sensor has the ability to capture infrared light but an infrared blocking inner filter protects the CCD sensor from high energy infrared light such as sunlight. This blocking filter was removed to allow the camera to capture that specific range of light spectrum. In addition, a filter to block visible light was added, to avoid undesired reflections on the cornea. To ensure the user's eye health it is important to take into account certain considerations. Any source of light radiation in large concentrations can cause damage to the human eye. The International Electrotechnical Commission (IEC) has established standards for human eye safety with various sources of light radiation. Document IEC-60825 deals with laser sources, and IEC-62471 deals with lamp sources and LEDs, under normal continuous or pulsed operation. This last document defines the maximum infrared radiation exposure (for wavelengths around 770 to 3,000 nm) allowed on the cornea for day-long continuous exposure. For this reason, the IR light source was set to deliver an irradiance of 3 mW/cm^2 , which is well below the IEC recommended maximum of 10 mW/cm^2 .

III. IMPLEMENTATION

Figure 3 shows the block diagram of the proposed system. The method is composed of four stages, which are described below:

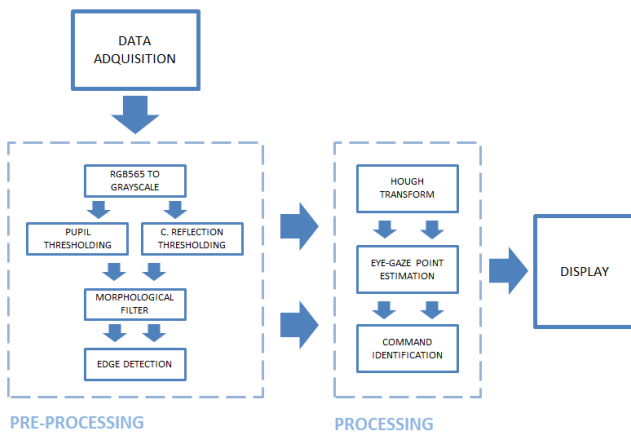


Fig. 3: Proposed method

1) Data acquisition:

Data acquisition is done through the video camera which is mounted on a support attached to the user's glasses, as shown in 4. This camera captures images from the user's eye and sends them to the processing unit. When the eye is illuminated by an infrared light positioned outside the optical axis, much of this light is absorbed by the pupil causing it to look darker. This effect is known as the "dark pupil effect", which improves the pupil-iris contrast, making easier the pupil segmentation.



Fig. 4: Eye tracking glasses

2) Pre-processing:

The pre-processing stage receives the data sent by the camera and conditions it for the next instance.

Although the acquired images are in RGB-565 format, the processing is performed only on the blue channel since it has the highest entropy. This data set is used as input of a bunch of algorithms that allow to identify the center of the pupil and the center of cornea reflection. The blue channel information is applied to two thresholding algorithms. Based on the dark pupil effect, the pupil thresholding seeks to obtain the darker pixels. On the other hand, the corneal reflection produces a region of pixels of high-intensity values caused by the reflection of the infrared light on the user's eye, as shown in Fig. 6, so the thresholding logic seeks to obtain the brighter pixels.

Then, an opening morphological filter is applied using a linear structuring element of dimension 1x3 [19], allowing to eliminate noise and isolated pixels due to previous thresholding.

Finally, edge detection is performed employing a 3x3 kernel Laplacian operator [13].

3) Pupil center and corneal reflections points:

The circular nature of the pupil and the reflection of the cornea cause the problem to be reduced to the search for centers of circumference through the HT. One of the great limitations of this transformation is that there is none feature analysis of the candidates to be centers of circumferences. Usually these are obtained only by voting in the parameter space, selecting the global maxima in the accumulation matrix [18]. The proposed method suggests some modifications that reduce the number of computation operations and improve the criterion of selection of circumference candidates.

Fig. 5 illustrates how the method works, where p_1 and p_2 are two arbitrary points belonging to the edge of the searched circumference. Assuming that such circumference has a radius R_o between R_{min} and R_{max} , the angular space is discretized in n points, and for each border pixel, instead of computing all the circumferences crossing over that pixel, we compute only those that have that pixel as center. The point p_3 , where the proposed circumferences intersect, is the point where the number of votes in the accumulation space is maximum, and matches with the center of the searched circumference.

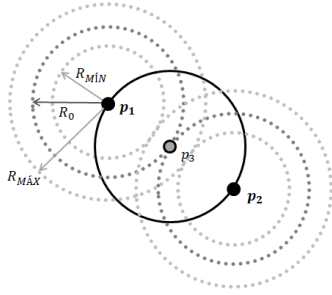


Fig. 5: Hough Transform implementation

The parametric space is obtained by Eq. 2.

Some imperfections in the image, in the edge detector, or simply the shape of the border under analysis, often produce regions that cause local or global maxima in the accumulator, despite not resemble at all to a circumference. To eliminate all these regions we propose to use the roundness shape features as a criterion of prior discrimination between regions that can produce local maxima, reducing in this way the number of candidates to be pupil centers. The candidate regions will be those whose roundness descriptor is approximately equal to 1.

4) Eye-gaze point estimation and command identification:

In eye tracking systems composed of a single camera and a single light source, if the user's head position remains fixed, relative to the camera position, it is possible to approximate the gaze direction, relative to the optical axis of the camera, computing the vector distance between the pupil centers and cornea reflection [22]. Thus, the gaze direction c_x, c_y is obtained through the distance vector between the pupil center c_{xpupil}, c_{ypupil} and the center of the cornea reflection c_{xIR}, c_{yIR} , according to the Eq. 3.

$$\begin{aligned} c_x &= c_{xpupil} - c_{xIR} \\ c_y &= c_{ypupil} - c_{yIR} \end{aligned} \quad (3)$$

It is known that the right human eye can rotate to the left up to 60° , to the right up to 90° , and upwards to 46° [23]. This means that the amplitude of eyeball movement is limited to the gaze direction. We use these known relationships as criterion to define the distance relationships between the center of the pupil and the center of the cornea reflexion to identify each movement command of the wheelchair: left, right and forward.

IV. RESULTS

To evaluate the performance of the method we used our own database, composed by two hundred pupil images acquired through the Discovery kit. These images belong to different people of different gender, age and eye color which were obtained under low, medium and high ambient lighting conditions. The performed tests include a total of 90 movement commands: 30 left, 30 right and 30 forward.

Environment Illumination	CHT Accuracy [%]	Proposed method Accuracy[%]
High	80	93.33
Average	58.62	86.2
Low	69.56	86.95

TABLE II: CHT and proposed method accuracy, tested on 200 images.

Fig. 6 (a) shows three captures taken under different ambient lighting conditions, (b) the performed edge detection and Hough candidates, (c) the center of circumferences obtained by the original CHT and (d) the center of circumferences obtained by the proposed method. In the very first two cases, the original CHT, performing the search by global maxima in the space accumulator, yields several points and is not be able to discern between each of these centers. On the other hand, the proposed method, based on the roundness factor, selects the point whose neighborhood is more like a circumference and at the same time have the highest accumulator value.

Table II shows the accuracy of the original CHT versus the proposed method under different ambient lighting conditions. The threshold values were adjusted empirically in order to maximize the original CHT accuracy. Targets were manually created, and it was considered as right detection, distances less than three pixels. Although this value is arbitrary, it takes into account the error of manually creating each target. Based on these results, the proposed method improves the original CHT response between 16% to 46% depending on lighting conditions.

Table III and Fig. 7 show the performance of the algorithm for each of the three possible movement commands of the wheelchair: left, right and forward.

The processing time of each video frame during the tests was about of 100 ms.

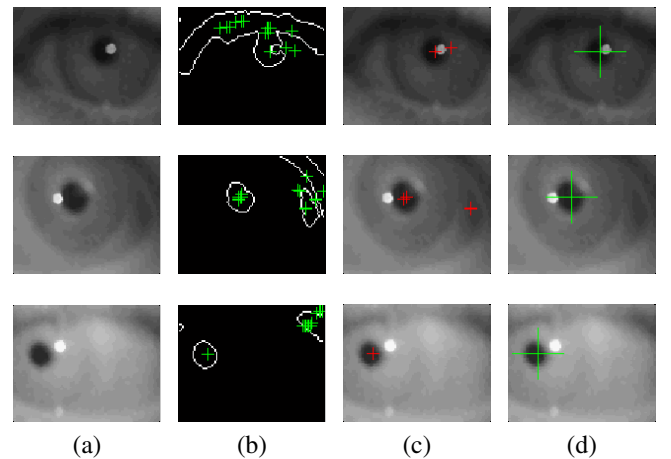


Fig. 6: Original CHT and proposed method performance under different ambient lighting conditions: (a) Original frame (b) Edge detection and CHT candidates (c) Pupil center by global maximum (d) Pupil center by proposed method

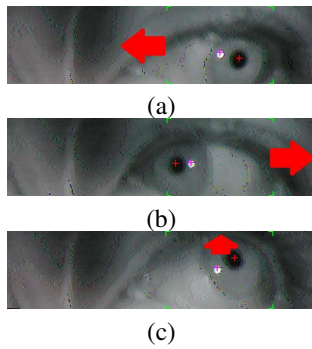


Fig. 7: (a) Subject eye pointing left (b) Subject eye pointing right (c) Subject eye pointing up

	COMMAND			
	LEFT[%]	RIGHT[%]	UP[%]	MEAN[%]
SUBJECT 1	92	89	80	87
SUBJECT 2	85	81	78	81
SUBJECT 3	90	89	83	87

TABLE III: Command identification accuracy, performed on 90 commands.

V. CONCLUSIONS

The proposed method allows real time gaze tracking using a very low cost computing platform, with satisfactory results under different ambient lighting conditions. The processing cost of each video frame allows to operate at a refresh rate of ten frames per second. Currently we are working on the implementation of the system on a commercial electric chair. As future work, in order to improve the system accuracy, we will research alternatives ways of selecting candidates for centers of circumferences of the CHT and the threshold value will be automated according to ambient illumination conditions. In addition, an algorithm to identify blinkers will be developed to reduce false detections and we will expand our eye database.

REFERENCES

- [1] J. D. Simeral, S. P. Kim, M.J. Black, J.P. Donohue and L.R. Hochberg, "Neural control of cursor trajectory and click by a human with tetraplegia 1000 days after implant of an intracortical microelectrode array", *J. Neural Eng.*, vol. 8, no. 2, April 2011.
- [2] Tsui, C.S.L., Pei Jia; Gan, J.Q., Huosheng Hu; Kui Yuan, "EMG-based hands-free wheelchair control with EOG attention shift detection, Robotics and Biomimetics", 2007. ROBIO 2007. IEEE International Conference on , vol., no., pp.1266,1271, 15-18 Dec. 2007. doi: 10.1109/ROBIO.2007.4522346
- [3] D.A. Craig; H.T. Nguyen, "Adaptive EEG Thought Pattern Classifier for Advanced Wheelchair Control", Proceedings of the 29th Annual International, Conference of the IEEE EMBS, Cité Internationale, Lyon, France, August 2007.
- [4] R. Berjon, M. Mateos, A. L. Barriuso, I. Muriel and G. Villarrubia, "Alternative human-machine interface system for powered wheelchairs", IEEE 1st International Conference on Digital Object Identifier, Serious Games and Applications for Health (SeGAH), 2011.
- [5] Umeshkumar Jaiswar; Sanjna S. Repal, "Real Time Breath Processing Based Embedded Vehicle Control Using WSN", International Journal of Science and Research (IJSR), 2013, ISSN (Online): 2319-7064
- [6] Gautham Krishnamurthy ; Maysam Ghovanloo, "Tongue Drive: A Tongue Operated Magnetic Sensor Based Wireless Assistive Technology for People with Severe Disabilities". Circuits and Systems, 2006. ISCAS 2006. Proceedings. 2006 IEEE International Symposium

- [7] Lund, M.E.; Christensen, H.V.; Caltenco, H.A.; Lontis, E.R.; Bentsen, B.; Struijk, J.J., "Inductive tongue control of powered wheelchairs", Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE , vol., no., pp.3361,3364, Aug. 31 2010-Sept. 4 2010. doi: 10.1109/IEMBS.2010.5627923
- [8] R. A. Cooper, M. L. Boninger, A. Kwarciak, and B. Ammer, "Development of power wheelchair chin-operated force-sensing joystick", Engineering in Medicine and Biology, 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society EMBS/BMES Conference, 2002. Proceedings of the Second Joint (Volume: 3)/23-26 Oct. 2002.
- [9] Päivi Majaranta , Andreas Bulling, "Advances in Physiological Computing", Eye Tracking and Eye-Based Human-Computer Interaction, 2014, ISBN 978-1-4471-6392-3, pp 39-65
- [10] Amer Al-Rahayfeh; Miad Faezipour, "Eye Tracking and Head Movement Detection: A State-of-Art Survey" IEEE Journal of Translational Engineering in Health and Medicine (Volume: 1), 2013, ISSN: 2168-2372
- [11] Cristian E. Ordoñez; Juan I. Pastore; Virginia Ballarin; Eduardo L. Blotta, "Real-Time Iris-Tracking Embedded System" SASE 2015, FI-UBA, Bs. As., Argentina, ISBN 978-987-45523-3-4
- [12] Pastore J.; Moler E., "Identificación de Senos Frontales mediante Factores de Forma y descriptores de Fourier", IEEE LATIN AMERICA TRANSACTIONS, VOL. 2, NO. 3, SEPTEMBER 2004 ISSN: 1548-0992, pp 157 - 161
- [13] P.M. Merlin, D.J. Farber, "A parallel mechanism for detecting curves in pictures", IEEE Trans. Comput., C-24, pp. 96-98, 1975.
- [14] D. H. Ballard, "Generalizing the Hough transform to detect arbitrary shapes", Pattern Recognition, vol 13(2), pp. 111-122, 1981.
- [15] P.V.C. Hough, "Method and means for recognizing complex patterns", U. S. Patent 3069654, 1962.
- [16] R.O. Duda, P.E. Hart, "Use of the Hough transform to detect lines and curves in pictures", Comm. Assoc. Computing, vol. 15, pp. 1-15, 1972.
- [17] P.M. Merlin, D.J. Farber, "A parallel mechanism for detecting curves in pictures", IEEE Trans. Comput., C-24, pp. 96-98, 1975.
- [18] J.S. Bruno, "Reconocimiento de burbujas en formularios inteligentes aplicando la Transformada de Hough", FRBA-UTN, Buenos Aires, Argentina, 2005.
- [19] Rafael C. Gonzalez y Richard E. Woods. "Digital Image Processing". Prentice Hall 2nd Edition 2002.
- [20] Daugman, John, "How iris recognition works", IEEE 2004.
- [21] Mulvey, F., Villanueva, A., Sliney, D., Lange, R., Cotmore, S., Donegan, "Exploration of safety issues in Eyetracking", Communication by Gaze Interaction (COGAIN) IST-2003-511598: Deliverable 5.4. Available at <http://www.cogain.org/results/reports/COGAIN-D5.4.pdf>
- [22] Elias Daniel Guestrin; Moshe Eizenman; "General Theory of Remote Gaze Estimation Using the Pupil Center and Corneal Reflections", IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING, VOL. 53, NO. 6, JUNE 2006
- [23] Spector, R.H.: "Visual fields" in Walker, H.K., Hall, W.D., and Hurst, J.W. (Eds): "Clinical methods: the history, physical, and laboratory examinations", Butterworths, Boston, 1990, 3rd edn.
- [24] STM32F4DISCOVERY Expansion Boards Datasheets, www.element14.com/stm32f4-expansion.
- [25] Miad Faezipour y Amer Al-Rahayfeh, "Eye Tracking and Head Movement Detection: A State-of-Art Survey". University of Bridgeport, Bridgeport, CT 06604, USA, 6 November 2013.
- [26] UBIRIS database, <http://iris.di.ubi.pt/>