# Controlling Hand-Assistive Devices: Utilizing Electrooculography as a Substitute for Vision

Yaoyao Hao, Marco Controzzi, *Student Member IEEE*, Christian Cipriani*, *Senior Member, IEEE,* Dejan B. Popović, *Member, IEEE,* Xin Yang, Weidong Chen, Xiaoxiang Zheng, and Maria Chiara Carrozza, *Member, IEEE*



The DOI of the final edited version of this paper is: 10.1109/MRA.2012.2229949

# Controlling Hand-Assistive Devices: Utilizing Electrooculography as a Substitute for Vision

Yaoyao Hao, Marco Controzzi, *Student Member IEEE*, Christian Cipriani*, *Senior Member, IEEE,* Dejan B. Popović, *Member, IEEE,* Xin Yang, Weidong Chen, Xiaoxiang Zheng, and Maria Chiara Carrozza, *Member, IEEE*

***Abstract*—A new application for electrooculography (EOG) based on the primary role of vision in human prehension function is here described: namely, the control of prehension of hand assistive devices (HADs) through visual features estimation of a target object. We hypothesized processed vertical and horizontal EOG signals while observing (border-scanning) an object, as a suitable method for estimating its visual features, which are essential for selecting the grasp affordance. To prove the hypothesis, firstly we recorded EOG signals in ten healthy subjects, while scanning different lines and objects. The measurements aimed at evaluating the successful recognition rate of five different shaped objects. Off-line analyses demonstrated that successful object recognition, was significantly high and ranged between 74.3%-97.1% across subjects. In order to assess the practical viability of the system we implemented it in real-time to control on-line a robotic hand in grasp tasks, and tested it on fifteen subjects. Outcomes showed the viability to differentiate between four affordances already after a short training.**

***Index Terms*—Assistive devices, control of prehension, electrooculography, human-machine interfaces.**

## I. Introduction

THE loss of hand function, due to amputation or neurological injuries, causes severe physical and psychosocial debilitation. The most evident and critical impairment after upper limb amputation or neurological injury like brachial plexus or spinal cord injury is the loss of the prehension i.e. the ability to perform those movements in which an object is seized and held partially or wholly within the compass of the hand.

There have been many attempts to build devices that replace or substitute the hand function to promote the autonomy of people with amputation or neurological injuries in recent years, and today *hand assistive devices* (HAD) are becoming the reality. These include hand prostheses [1], robot hands mounted on manipulation aids (like robot arms on wheelchairs), assistive hand-exoskeletons [2] and electrical stimulation technology for the training of the grasping by individuals with hand dysfunctions [3]. What is still lacking is a robust, reliable, and intuitive control signal and interface, allowing a large bandwidth (bit rate) communication channel to control available assistive systems.

To this aim human-machine interfaces (HMI) have so far been extensively studied for people who, due to the pathology, have no voluntary control of their upper limbs or have no limbs at all. Cortical signals recorded invasively directly from the primary motor cortex or biopotentials recorded through surface electroencephalography, have been used to control robotic arms, hand orthoses, and prostheses [4]-[5], with the aim of restoring the connection from the brain to a paralyzed arm. Complementarily for hand amputees, although invasive techniques have shown to be viable physiologic means for controlling dexterous prostheses, low bit rate, non-invasive interfaces such as surface electromyography (EMG) from the residual limb have been the most employed solution so far [6]. Despite these efforts, the main challenge in the assistive engineering/robotics field is still today in the use of control signals and extraction of intent, and electrooculography (EOG) represents one of the possible solutions.

EOG is a non-invasive recording technique that allows the resting corneoretinal potential of the eye to be recorded through electrodes properly positioned. The physiological basis of such potential – namely the electrooculogram – is that the cornea is electrically positive relative to the back of the eye, and this source behaves as if it were a single dipole oriented from the retina to the cornea. Since eye movements produce a rotating dipole source, accordingly, electrooculograms are an effective measure of the horizontal and vertical rotations. While the potential had been discovered much earlier (by Emil du Bois-Reymond in 1848) and the first pioneering clinical application was developed in the 1960's by Arden *et al.*, [7], the use of EOG in the clinical and research laboratory came of age with the advent of stable electronic signal conditioning

Manuscript received 15 March, 2012. This work was supported by the European Commission, under the WAY Project (EU-FP7-ICT-288551), and by grants from the National Natural Science Foundation of China (No. 60873125, 3080028, 61031002, 61001172), the Zhejiang Provincial Natural Science Foundation of China (No. Y2090707) and the Ministry of Science and Technological Development of Serbia (research grant #175016). *Asterisk indicates corresponding author.*

*C. Cipriani, M. Controzzi and M. C. Carrozza are with The BioRobotics Institute, Scuola Superiore Sant'Anna, 56025 Pontedera, Italy (e-mail: ch.cipriani@sssup.it).

Y. Hao and X. Zheng are with the Qiushi Academy for Advanced Studies, the Department of Biomedical Engineering, and the Key Laboratory of Biomedical Engineering of Ministry of Education, Zhejiang University, Hangzhou 310027, China.

D. B. Popović is with Faculty of Electrical Engineering, University of Belgrade, 11000 Belgrade, Serbia and Center for Sensory-Motor Interaction, Aalborg University, 9220, Aalborg, Denmark.

W. Chen and X. Yang is with the Qiushi Academy for Advanced Studies, and the College of Computer Science, Zhejiang University, Hangzhou 310027, China.

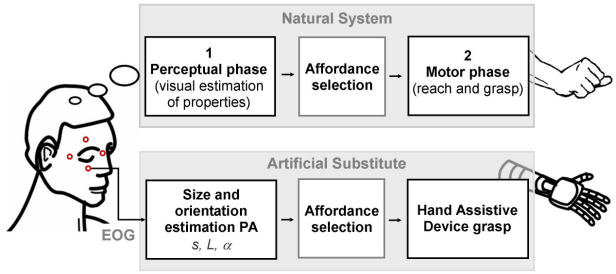



Fig. 1 Natural prehension control system vs. artificial EOG-based substitute. The biological grasp is composed of a perceptual phase where the properties of the object are visually estimated and the proper affordance (how to grasp the object) is selected; this is then executed by the motor system in the second phase. In this artificial system, the perceptual phase is replaced by a processing algorithm (PA) applied on EOG signals while scanning the target object, able to recognize shape (sides length: *s* and *L*); such information may be used to drive the preshaping phase, hence the grasp of a hand assistive device. The orientation ($\alpha$) estimation was not a part of this study.

amplifiers and filters. Since the 1980's EOG has been widely exploited as a control channel in a number of communication or assistive devices for the severely disabled patients [8]-[9]. Other eye movement recording methods include infrared reflection devices, scleral search coil, and video-oculography; advantages and disadvantages of each of these methods are detailed in [10].

This work tries to bring a new dimension in the control of hand assistive devices. The novelty follows motor control studies that emphasize the role of vision in the reaching and grasping function. In humans the selection of how to grasp (the so called *affordance*) depends not only on the function that has to be achieved but also on the estimated properties of the object and is the result of a complex set of perceptual to motor transformations [11]. Properties like distance and orientation in the environment are required to transport the hand in the correct orientation, direction and for a particular distance; however, other properties like size and shape are necessary to control the grip aperture and to select the most appropriate grasp points [12]. The action of grasp can hence be roughly divided in a first phase where a perceptual (imagined) estimation of the properties of the object is derived in order to select the proper affordance, which is in turn implemented in the second phase by the motor system (action). This analysis of the prehension is at the basis of the visuo-motor channel hypothesis proposed by Jeannerod [12], and is sketched in the upper part of Fig. 1.

Based on the primary role played by natural vision in reaching and grasping, recently Došen *et al.*, developed an artificial vision system based on a web camera and distance sensors acting as a substitute for the perceptual element in grasping [13]. Their system used the measured distance to the target object and computer vision algorithms to estimate the size and orientation of the object and to output the corresponding affordance commands for the control of prehension. More recently they demonstrated the real-time feasibility of such system on HADs [14].

Our research takes inspiration from the importance of vision in planning movement and merges it with the studies on physiologic/neural control, by presenting and demonstrating the feasibility of a new application for EOG not yet investigated in literature, i.e. the control of prehension of a hand assistive device by means of visual features estimation of a target object. In particular we hypothesized approximating objects to rectangles and processing the vertical and horizontal EOG signals while scanning the object, as a method for estimating its size, which provides key information to select the grasp affordance, hence for grasping (Fig. 1). In fact, the first phase of a grasp (i.e. the *preshaping*), consists in orienting and opening the hand so that the object fits comfortably.

Within this paradigm, an artificial substitute replaces the natural perceptual estimation of the properties of an object and by recognizing it, selects the correct motor program (affordance) (Fig. 1). We directly derive the control signal for preshaping from natural vision (i.e. from the scanning of an object). The suggested method would rely on user ability and since it is based on object observation holds the promise to become an intuitive and flexible HMI, once learning by the user has taken place.

In order to prove the concept of such EOG application, firstly, ten healthy subjects were enrolled in three preliminary experimental tasks for assessing (i) horizontal and (ii) vertical resolution and (iii) object recognition. The goal of the first two was to identify the sensitivity of the EOG signal by subjects while looking at lines with different lengths, in order to infer on the minimum recognizable differences between two objects. The third experiment was aimed at evaluating –with off-line processing- the successful recognition rate of different standing objects.

After these preliminary and necessary steps, the off-line algorithms were ported onto a real time system which was tested by a group (different from the previous one) composed of fifteen healthy subjects. Specifically, using the EOG on-line system and a programmable robotic hand (acting as a HAD), they performed *grasp* and *release* tasks with objects requiring four different grasp affordances, within experimental protocols aimed to provide insights on the robustness of the concept and on short-time training effects.

This paper presents the following: (i) the concept, (ii) processing algorithm for identifying the object visual features, (iii) experimental setups, and (iv) the results achieved both in off-line and on-line experiments.

## II. Methods

The prehension control system consists of three main modules: the EOG signal acquisition unit, the processing algorithm (PA) and the hand assistive device (cf. Fig. 1).

### A. *EOG Signal Acquisition*

The feasibility of the application was evaluated in experimental trials whereby subjects' EOG were recorded while "scanning" (see definition below) different objects or lines. Four-channel EOG signals were acquired with the Synamps RT amplifier (Neuroscan, USA) using active Ag/AgCl electrodes, referenced to an electrode placed on left earlobe, and grounded at another electrode placed on right earlobe. Two electrodes were placed above and below the right eye and two to the left and right of the eyes according to a bi-temporal electrode application [9]. EOG signals were digitized at a sampling rate of 2048 Hz and stored for off-line analysis.

### B. *Processing Algorithm*

Since EOG is proportional to the rotational angle of the eye, the fixation point can be detected from the recorded signal when the distance between eyes and objects is fixed and the patterns of potentials can be used to resolve the reverse task, i.e., that of describing the gaze trajectory, such that the **eye following the borders of an object clockwise or anticlockwise** (here defined *border-scan*) can be used to estimate the object shape.

This concept is summarized in Fig. 2 (a): differential signals from right-left and above-below electrodes form the horizontal (HEOG) and vertical (VEOG) EOG channels, respectively. These are digitally filtered with a 4th order Butterworth low-pass filter (cut-off frequency at 20 Hz) to reduce interferences from power-line, high frequency noises and other physiological sources (e.g. EMG artefacts) and subsequently used to estimate the gaze trajectory when the subject border-scans the object. As an example, the shape of a parallelogram having a 30 degrees orientation (of the longer axis with respect to the horizontal axis of the image plane) [cf. Fig. 2 (b)] is described. This example was selected since most objects that can be held with one hand (and are used in daily life) can be reduced to the form of parallelepipeds, and their projection in the 2-dimensions plane (i.e. the vision representation) are parallelograms, or even rectangle like. Fig. 2 (c) shows on an x-y plot (rescaled in millimetres) the HEOG versus VEOG channels recorded while anticlockwise scanning the parallelogram starting from the lowest point (the bottom corner), and keeping head still (a moving head would affect the eye gaze). Under these hypotheses, the x-y signal representation transforms into a *pseudo-parallelogram*, in which eight key data are of interest: e.g. either the position of the four corners (A i.e. the starting point, B, C, D) or the equations of the four sides (AB, BC, CD, DA). Focusing on the corners, once they are identified, in theory the slope of the side AB (or CD) α (and similarly θ) can be computed through the trigonometric equation:

$$tg\alpha = \frac{VEOG_B - VEOG_A}{HEOG_B - HEOG_A} \qquad (1)$$

where the numerator and denominator of the fraction are the

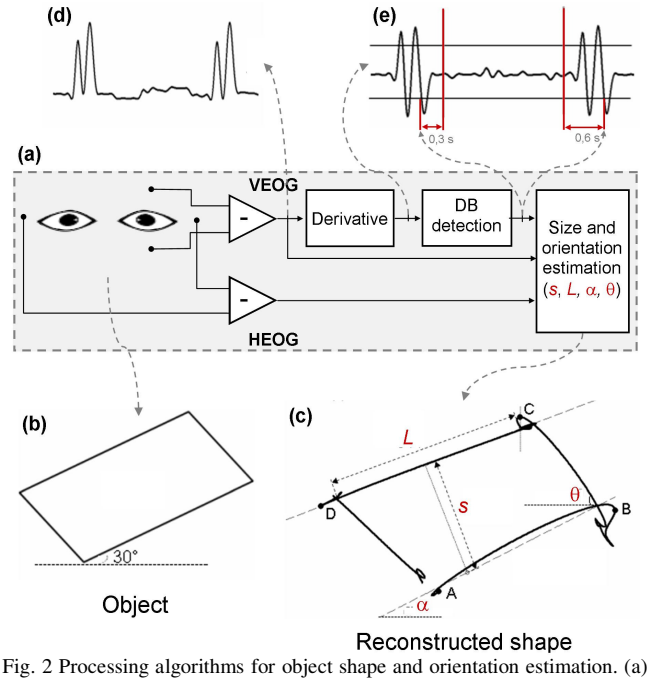



Fig. 2 Processing algorithms for object shape and orientation estimation. (a) EOG processing block diagram. (b) 30 degrees oriented parallelogram and its reconstructed version using EOG signals recorded while border-scanning it in (c). The long and short sides (*L, s*) and angles (α, θ) of the object are estimated from the reconstructed shape. (d) VEOG signal with two double-blinks (DB) and (e) its differential result. The horizontal lines mark the upper and lower thresholds which are used to detect double-blinks; the vertical lines mark 300 ms after the first DB and 600 ms before the second DB, i.e. δενοτε τηε τιμε ωινδοω ιν ωηιχη τηε ΕΟΓ σιγναλ ισ υσαβλε φορ ϖι συαλ φεατυρεσ εστιματιον. Τηε οριεντατιον (α) estimation was not a

horizontal and vertical gaze-shifts. With practical EOG though, a better estimation can be obtained mathematically through a straight line fitting (least squares fitting) from A to B and extracting its slope. After this the long side of the object *L* and similarly the short side *s* could be theoretically calculated using the Pythagorean theorem.

Nevertheless, as this was a preliminary study, we simplified the problem and just considered objects standing on one side, (i.e. with an orientation α equal to 90 degrees), and decided to approximate them as rectangles (α=90°−θ). Hence, equations become easier, and in the case of a rectangle standing on the short side (side DA), *L* and *s* are:

$$\begin{aligned} L &= c_v \times \left(\max(VEOG) - \min(VEOG)\right) \\ s &= c_h \times \left(\max(HEOG) - \min(HEOG)\right) \end{aligned} \qquad (2)$$

where $c_v$ and $c_h$ are vertical and horizontal mapping coefficients which are required to transform amplitude of EOG (uV) to effective dimensions of the object (mm), respectively (such coefficients can be retrieved after an initial calibration with a given rectangle at a given distance). Generally, if the starting point is different from the bottom corner (e.g. B instead of A), or it belongs to one side, or if the x-y signal badly resembles the object, the four corners could still be found, and its shape and orientation still be retrieved through signal processing (e.g. filtering) and simple numerical methods (e.g. best fitting of a parameterized parallelogram).

Within this paradigm, the beginning and end of a *border*

*scan* can be effectively marked by the user in the vertical EOG signal through two successive eye blinks (hereafter double blinks, DB), and the HEOG and VEOG traces between two DB can be exploited to estimate object visual features as described above. In order to detect voluntary eye double blinks, the VEOG is differentiated. An example is reported in Fig. 2 (d) and (e), showing the VEOG signal and its corresponding differential version. Once differentiated, the DC component of the signal is cancelled out, and the DB becomes easily recognized through a finite state machine (FSM) implementing a threshold method: if the differential signal exceeds the upper and lower thresholds in a predefined sequence and duration, a DB is detected. However, as also depicted in Fig. 2 (e), since the VEOG signal is in transient state around the DB, signals should be discarded and not used for object estimation for a certain time window: pilot experiments demonstrated that 300 ms after the first and 600 ms before the second DB are suitable values, also across subjects.

Resuming, the shape and orientation of an object can be approximately reconstructed by border-scanning it, and by preceding and following such scan with a DB. In such paradigm the DBs are employed to mark the initial and final instants, whereas the VEOG and HEOG signals track the rotational angles of the eye, therefore allowing determination of the object features. If the workspace is fixed also the dimensions of the object may be estimated, as the gaze shifts are related to the size of and distance to the object. The reconstructed size and orientation of the object can be functionally used to control the preshaping phase of a grasp, therefore prehension of a HAD.

### C. Experimental Protocol (Off-line EOG)

Ten able-bodied young subjects (8 males and 2 females whose average age was 29.8 ± 3.9 years old) who claimed to have normal or corrected vision were enrolled in this study. Nine were naive to EOG, and one (subject no.1) was a participant with significant experience in the use of EOG systems. Specifically, this participant had put in place the experimental set up and for this reason had trained for a period quantified in 10 days, 45 minutes each day. One by one, the subjects were seated on a chair in front of the target which was located 71 cm from the line of their eyes on a desk. This distance was chosen as it is coherent with the *functional reach distance* in humans. At the beginning of the experiments, each subject briefly trained: he/she was asked to scan by sight, calibration rectangles on a PC screen at fixed distance (71 cm), anticipated and followed by DBs in order to gain confidence with the system and learn how to perform DBs. This preliminary procedure was also necessary to adjust the thresholds for the DB detection, and lasted for about 5-10 minutes. After that, three different experiments were carried out in a sequence: (i) horizontal and (ii) vertical resolution assessment, and (iii) object recognition. Although it is known from literature that with humans and primates smooth pursuits tracking (i.e. slow shift eye movements) is more efficient in the horizontal than in the vertical dimension [15], the aim of the first two experiments was to: (i) acquire knowledge on the horizontal and vertical sensitivity of the EOG scan using the recording system available and simple Equations (2), in order to infer on the minimum recognizable differences between two objects, and to (ii) further train subjects before performing the following test. The latter, aimed at evaluating the successful recognition rate of different objects. The experimental setup is illustrated in Fig. 3; the first two experiments were carried out with the targets (i.e., the lines) in position *P1*, the last experiment with the targets (i.e., the objects) in position *P2*. It is important to note that the experiments were carried out in a life-like scenario: (i) no instructions on the speed of the eye-scans were given (volunteers could decide based on their confidence); (ii) the heads of the participants were free to move (although they were asked to keep them fixed); (iii) other people were working/moving around in the same room; (iv) objects were randomly coloured and shaped; (v) on the desk there were other objects, and finally, (vi) no special background behind the objects was used.

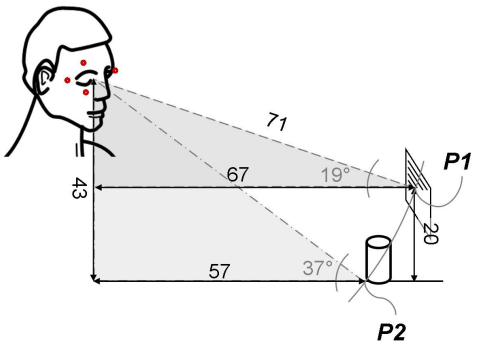



Fig. 3 Experimental set-up. The vertical and horizontal resolution assessment tests were carried out with the lines in position P1; the object recognition test was performed instead, with the objects in position P2 (eye elevation of 37 degrees). In both cases the distance was about 71 cm.

TABLE I
LIST OF OBJECTS USED FOR THE OFFLINE EXPERIMENTS

| Object | Size ($s \times L$) [mm] | Possible grasp |
|---|---|---|
| Cylinder | 70 × 120 | Cylindrical |
| ½ liter bottle | 60 × 190 | Cylindrical |
| Felt tip pen | 25 × 125 | Tripod/lateral |
| Seal tape | 95 × 50 | Spherical |
| Small ball | 38 × 38 | Pinch/Tripod |

#### 1) Horizontal and vertical resolution assessment

Subjects were asked to scan by sight -back and forth- horizontal lines employing the DB marking method described above. Five thin lines drawn on a vertical plane at a (mean) distance of 71 cm (position P1 in Fig. 3) from the subject's right eye were tested 10 times each. The length of the five lines was 100, 105, 110, 115 and 120 mm, therefore the differences among them ranged between 5 and 20 mm. The eye rotation angles required to scan the lines were in the range of 8-9.5 degrees. The experiment was subsequently replicated with the lines oriented vertically. The length of the lines was estimated off-line using Equations (2); the Spearman's rank correlation coefficient was used as the performance metric. Indeed, the latter assesses how well the relationship between two variables

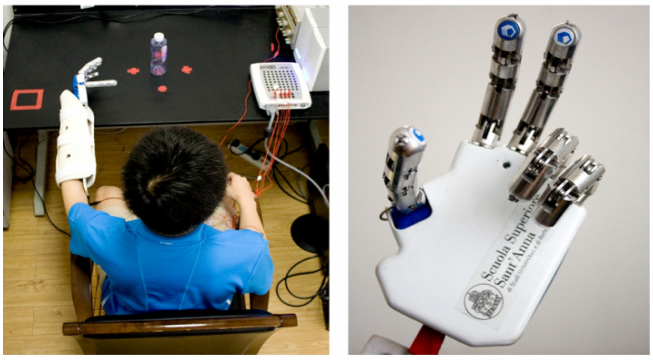
(in this case effective line lengths and estimated line lengths) can be described using a monotonic function.

*2) Object recognition*

Maintaining the distances illustrated in Fig. 3 (this time in position P2), subjects were asked to scan five common objects standing on their bases, having various dimensions (cf. Table I) and representing the power, precision and lateral grasps used in activities of daily living. All the objects were presented to the viewer vertically (orientation of the longer axis). Size estimation ($L$ and $s$) was performed off-line, implementing Equations (2) on the stored EOG signals. The mean errors on the estimated sides (or axes) for each object were used as performance metrics.

### D. On-line Experiments with Physical HAD

After the preliminary experiments (described above), the off-line algorithms were ported onto a real time system which was tested on-line by a new group of subjects. Fifteen healthy participants (12 male and 3 female, aged 24.7±1.8) took part in a set of experiments aimed to evaluate the practical viability and real time performance of the system proposed. Specifically, using: (i) the EOG electrodes and acquisition system; (ii) a personal computer running in real time equations (2) and an affordance-selection algorithm; (iii) a "physical" artificial hand (in place of a HAD); and (iv) a grasp-triggering system. Subjects carried out *reach, pick* and *place* tasks. What was really important here was to assess the ability of using EOG to select the correct affordance, rather than the task itself. However, the *reach, pick* and *place* task with the robotic hand, in fact seeming like a game, helped keep the participants focused.



(a) (b)

Fig. 4 Set up used for the on-line experiments with the physical HAD. (a) Participant wearing the orthopedic splint attached to the hand. The red marks on the table denote the positions of the workspace. (b) Robotic hand used for the real time experiment emulating a *hand assistive device*.

*1) Real-Time Affordance-Selection Algorithm.*

A real-time algorithm for selecting the affordance appropriate to an object, was developed based on the work by Došen et al., [14]. In particular, the algorithm was capable of recognizing two DB, of estimating the sizes ($L$ and $s$) of an object scanned [using Equation (2)], and –based on these– of selecting the required preshaping posture for a robotic hand according to a reduced version of the IF-THEN rule proposed by Došen et al. [14]. The estimated object size was compared to the pre-defined threshold values and one out of four preshaping postures (cylindrical, lateral, tripod or bi-digital) was selected. Specifically, if $L$ was greater than a fixed threshold $L^*$ (i.e. *long object*), cylinder or lateral grasp was selected; alternatively (a *short object*), a tri-digit or bi-digit grasp was chosen. If $L$ was greater than $L^*$ (*long object*), and if $s$ was greater than $s_1$, a cylindrical grasp was selected; otherwise, a lateral grasp. For a short object, if $s$ was below $s_2$, a bi-digit grasp was selected, alternatively a tri-digit grasp was the choice. The algorithm was able to select one of the four affordances and send a command to the hand within 1 ms after the second DB.

*2) HAD: Artificial Hand*

The robotic hand employed, was a commercial version of the SmartHand [1] (commercialized by Prensilia – Italy), shown in Fig. 4(b); this was able to perform the four grasps of interest for this study. The hand was fixed to an orthopedic splint, that in turn was worn on subjects' left forearm so that they could move it, as shown in Fig. 4(a). Since it was available a robotic hand was used, but an assistive hand-exoskeleton could also have been used.

*3) Grasp-Triggering System*

The triggering system consisted of a single-channel, myoelectric electrode and acquisition apparatus that collected surface EMG signals from muscles in the anterior compartment of the left forearm, and was able to recognize (wrist) flexion contractions with a minimum delay (40 ms), through a simple threshold technique. It was used as a means for voluntarily closing the fingers, starting from their current preshaping posture (selected through the affordance-selection algorithm), and for reopening it, to release the object at the end of the task. The experimental task is detailed hereafter.

*4) Experimental Protocols (On-line EOG)*

Participants sat on a chair in front of the target object which was placed in a specific location within a workspace, that was centered [cf. Fig. 4(a)] 71 cm far from their eyes. Eight different objects (cf. Table II: two for each affordance allowed by the hand), were used as targets, and were presented in a random order, in two consecutive blocks for each session (16 trials). Participants were instructed to operate the EOG system on the object (i.e. DB – border-scan – DB), which in turn would make the hand preshape in one of the four postures; if the preshaping was incorrect, they were allowed to repeat the border-scanning procedure until the correct preshaping was achieved (visual feedback). Once achieved, the participant could physically reach the object, trigger the grasp by means of a wrist flexion, pick the object and transport it to a *home-position* for release [cf. red square in Fig. 4(a)].

Fifteen participants (S1-S15) performed the experiment in one single session, with the eight objects placed in the centre of the workspace, for a total of 16 trials. To get insights of the learning of the procedure five of the participants (S1-S5) repeated the session twice within the following two weeks (*repeated test*). In order to evaluate the system robustness, five other participants (S6-S10), performed the experiment with four objects (O1,O3,O4,O7), each one placed on six different positions: i.e. the centre of the workspace, and 5 points 10 cm far from it on the front, the back, left, right and above [see red

marks in Fig. 4(a)]. This *workspace test* was done twice for two consecutive days (48 trials per subject each day).

In all cases, subjects were asked to keep their head as still as possible during the scanning process, but no physical systems were used to block the head. Before the very first session each participant, received instructions on how to complete the task and was allowed to freely train 10 minutes in order to learn: (i) how to perform a correct DB and border scan procedure; (ii) how to operate the triggering system to grasp/release objects; and (iii) how to operate the artificial hand to functionally grasp objects. After such short training subjects were asked to scan a standard rectangle (sized 148 mm × 105 mm) to calibrate the system, then the experiments started. Since timing was highly variable among participants, the number of border-scans required to select the correct affordance was used as the on-line performance metric.

TABLE II
LIST OF OBJECTS TARGETED FOR THE ON-LINE EXPERIMENTS

| Object ID | Description | Size ($s \times L$) [mm] | Supposed grasp |
|---|---|---|---|
| O1 | Shampoo bottle | 25 × 126 | Lateral |
| O2 | Felt Tip Pen | 13 × 180 | Lateral |
| O3 | ½ litre bottle | 56 × 152 | Cylindrical |
| O4 | Tea box | 94 × 136 | Cylindrical |
| O5 | Candy box | 55 × 55 | Tripod |
| O6 | Small ball | 62 × 62 | Tripod |
| O7 | Eye drops bottle | 26 × 53 | Bi-digital |
| O8 | Rubber | 22 × 33 | Bi-digital |

Dimension *L*: long axis, *s*: short axis

TABLE III
HORIZONTAL AND VERTICAL RESOLUTION TEST RESULTS

| | | S1* | S2 | S3 | S4 | S5 | S6 | S7 | S8 | S9 | S10 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| H | ρ | 0.716 | 0.681 | 0.878 | 0.758 | 0.865 | 0.530 | 0.756 | 0.729 | 0.606 | 0.936 |
| | p | $< 10^{-8}$ | $< 10^{-7}$ | $< 10^{-12}$ | $< 10^{-19}$ | $< 10^{-15}$ | $< 10^{-4}$ | $< 10^{-9}$ | $< 10^{-8}$ | $< 10^{-5}$ | $< 10^{-21}$ |
| | t | 1.43 | 0.97 | 2.57 | 1.15 | 1.03 | 2.32 | 2.28 | 1.21 | 1.39 | 1.21 |
| V | ρ | 0.699 | 0.563 | 0.019 | 0.558 | 0.359 | 0.241 | 0.217 | 0.107 | 0.440 | 0.491 |
| | p | $< 10^{-7}$ | $< 10^{-4}$ | 0.921 | $< 10^{-4}$ | 0.01 | 0.092 | 0.131 | 0.406 | 0.001 | $< 10^{-3}$ |
| | t | 1.48 | 1.23 | 2.46 | 1.44 | 1.38 | 2.33 | 1.86 | 1.35 | 1.56 | 1.48 |

Legend: H, horizontal resolution test; V, vertical resolution test; ρ, Spearman's rank coefficient correlation; p, p-value; t, mean time to execute the line-scan. Grey backgrounds highlight those tests for which ρ was not statistically significant (i.e. $p > 0.05$). The asterisk denotes the experienced participant.

TABLE IV
MEAN ERRORS ON THE ESTIMATED SIZES DURING OBJECT RECOGNITION TESTS AND MEAN BORDER-SCAN DURATIONS FOR ALL SUBJECTS

| Object | Object size [mm] | Estimated mean size (mean ± st. dev.) [mm] | H mean err [mm] (relative error) | V mean err [mm] (relative error) | Mean error (Euclidean distance) [mm] | Mean border-scan duration [s] |
|---|---|---|---|---|---|---|
| Cylinder | 70 × 120 | 70.8±9.6 × 125.3±24.3 | 0.8 (1.2 %) | 5.3 (4.4 %) | 5.4 | 2.94 |
| ½ liter bottle | 60 × 190 | 54.3±7.7 × 184.1±28.0 | 5.7 (9.5 %) | 5.9 (3.1 %) | 8.2 | 2.33 |
| Felt tip pen | 25 × 125 | 32.7±8.9 × 121.4±21.4 | 7.7 (30.8 %) | 3.6 (29 %) | 8.5 | 2.52 |
| Seal tape | 95 × 50 | 89.9±12.2 × 72.9±22.6 | 5.1 (5.4 %) | 22.9 (45.8 %) | 23.5 | 2.90 |
| Small ball | 38 × 38 | 32.2±9.1 × 54.6±21.3 | 5.8 (15.2 %) | 16.6 (43.7 %) | 17.6 | 2.05 |

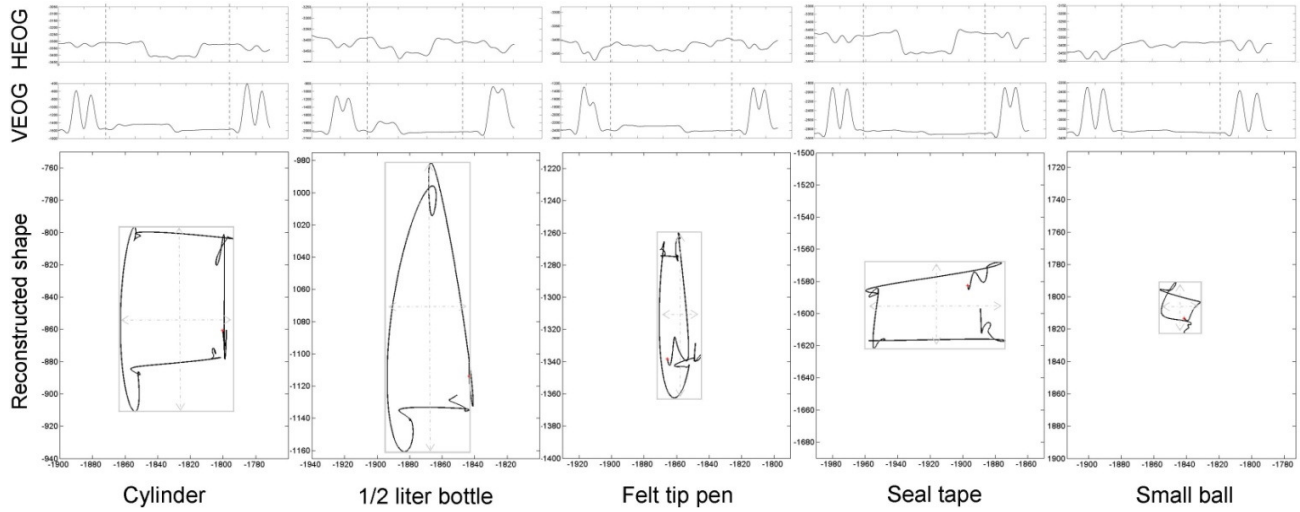



Fig. 5 Representative EOG signals and reconstructed shapes from subject no. 9 while border-scanning the five experimented objects. The upper row shows HEOG and VEOG time plots (vertical and horizontal axes units are millivolts and seconds respectively); the vertical dashed lines indicate 300 ms after the first DB and 600 ms before the second DB, and within the resulting time window the EOG signal is used for visual features estimation. The lower row demonstrate the reconstructed shape of the objects obtained by simply x-y plotting -after rescaling- the corresponding VEOG and HEOG signals. Superimposed on the graphs the grey rectangles denote how the processing algorithm estimates the size using Equation 2.

## III. RESULTS

### A. *Off-line EOG Experiments*

The ten subjects enrolled in the preliminary part of this work performed 150 *border-scans* each (5 horizontal and 5 vertical lines, 5 objects), for a total of 1500 recordings. The duration of each experimental session (each subject) was about 30 minutes (from electrodes placement to the end): each of the tests lasted around 10 minutes.

#### *1) Horizontal and Vertical Resolution Assessment*

The mean time to perform back and forth a line-scan (within the two DB) was subjective but relatively fast: 1.63 seconds on the average (i.e. a mean speed of 11 degrees/second). Hence experimental results here described are based on realistic conditions.

Table III shows the Spearman's rank correlation coefficient ρ, the p-value and the mean time to perform a back and forth line scan *t,* for each subject involved in the tests. As regards ρ, the closer it is to 1, the better effective and estimated lengths can be described by a monotonic function, whereas the p-value tests whether the observed ρ is significantly different from zero (a level of $p<0.05$ was selected as the threshold for statistical significance). Within the tested experimental conditions results confirm previous findings (as those in [15]): they are highly subjective and as denoted by the correlation coefficients, horizontal sensitivity is always better than the vertical one. In the horizontal test the best performer (subject no. 10) achieved a superlative correlation coefficient of 0.93, whereas the worse (subject no. 6) the modest value of 0.53.

#### *2) Object Recognition*

Representative results from the object recognition tests are presented in Fig. 5: time plots demonstrate the border scan signals resembling the five experimented objects by one of the subjects. Signals are those stored during the tests (sampling rate of 2048 Hz, after the 4$^{th}$ order digital filter, as described in the methods section), with no further smoothening or off-line processing. The graphs also depict the representative estimated lengths of the short and long axes of the objects. In these cases, the reconstructed pseudo-rectangles are somehow neat and the original shapes can be correctly estimated also at a glance; however, in other cases (not shown) the signals are not as good, and may yield to a worse estimation. The graphs in the lower row also denote the main drawback of the algorithm used; by comparing the EOG-based pseudo-rectangles (black curves) with the results of Equation (2) (grey rectangles) it is easy to understand that the latter is not able to separate the pure object shape from the EOG overshoots caused by the fast movement of the eye (especially marked on the VEOG). This results in an overestimation of the lengths of the object axis, but since this is likely to be a systematic error (depending on the subject), it was accounted when converting EOG recording units (from millivolts) to effective lengths (millimetres).

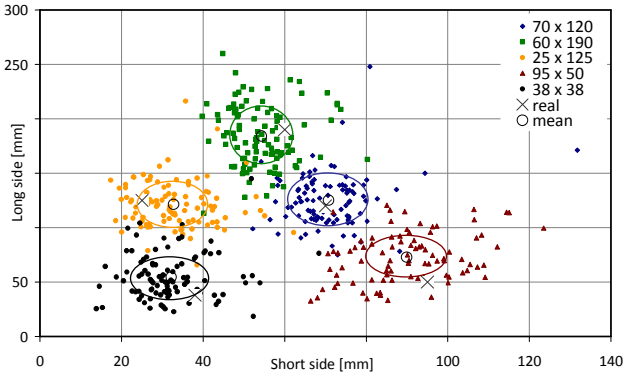



Fig. 6 Scatter plot in millimetres of the reconstructed horizontal and vertical dimensions of the objects, i.e. sides s and L respectively. The ellipses semi-axes sizes denote the standard deviation (on vertical and horizontal axes) of each sample group. Data from all ten subjects and five objects is included (500 samples, 100 for each object).

The scatter plot in Fig. 6 presents on an x-y plane the estimated dimensions for each border-scan to all objects by all subjects. The data (500 pairs of estimations in total) is displayed as a collection of points, each having the value of the estimated short side determining the position on the horizontal axis and the value of the estimated long side determining the position on the vertical axis. Samples related to the same object are coloured in the same way and black circles and crosses mark the mean estimation and the effective size for each object, respectively. The ellipse horizontal and vertical semi-axis lengths are sized as the standard deviations of the estimated short and long sides respectively, for each object. To shift from EOG recording units (volts) to millimetres, each subject's measurement is rescaled using the mean ratio from his/her recordings. The graph reveals a clear clustering among the five clouds, therefore significant separation among the five experimented objects. To confirm this statement a k-nearest neighbours (*k-nn* with k equal to 1) algorithm to recognize the five objects was applied on each subject's data as follows: the first three estimation pairs (estimated *L* and *s*) for each object were used to train the *k-nn*, whereas the last seven pairs were used for evaluation. Results are very good (86.2% on average) and range between 97.1% (best object recognition rate, for subject no. 9) to 74.3 % (worst result, subject no. 6). If the first five estimation pairs for each object are used to train the *k-nn* and the last five are used for evaluation (i.e. a slightly longer training procedure), the object recognition rate reach 100% success for all subjects.

Table III resumes the mean size estimations together with the mean horizontal, vertical and Euclidean distance errors. Coherently with the experimental horizontal and vertical resolution results, absolute errors on the horizontal signal are generally lower than those on the vertical axis. Only for the 25x125 mm felt-tip pen the horizontal error is greater than the vertical one, and in just two cases the errors (in particular the vertical errors) are greater than 10 mm. Such cases refer to those objects having shorter long sides (i.e. 38 and 50 mm).

Table IV also resumes the mean times to perform the border-scan for the five objects; these were generally fast, ranging between 2.05 (for the small ball) to 2.94 seconds (for

the cylinder).

### B. On-line Experiments with Physical HAD

Fig. 7 shows the distribution (in percents) of number of scans required to achieve correct preshaping, by the fifteen naïve participants, after one single session. It shows that overall, across subjects and objects, 65% of trials were successfully achieved with one border-scan; in 15% of cases two scans were required. Hence, around 80% of trials were achieved with one or two border scans. We believe this was an interesting result for participants naïve to the system.

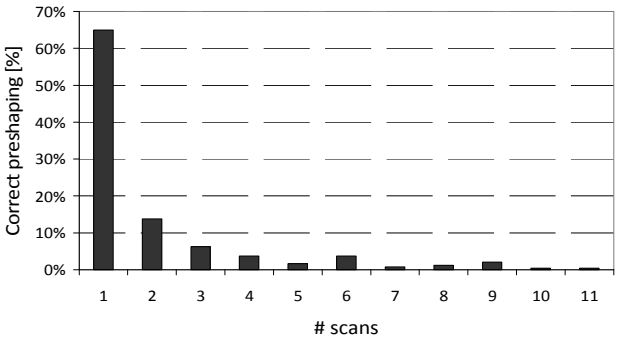



Fig. 7 Distribution of number of border scans required by 15 healthy subjects to generate correct grasp affordance, through the EOG on-line system.

Overall, subjects were able to select the proper affordance after 1.9±1.9 (mean ± st. deviation) border scans. The high standard deviation denotes large variability based on subjective performance, and the result is likely to be due to the short training. Although not supported by statistics, the training curve across the three-sessions/days repeated test (cf. Fig. 8), shows a positive trend (*p-value = 0.09*); with some practice the number of scans required diminished (from 2.2±0.8 in day1-block 1 to 1.5±0.2 in day3-block 2). The distribution across the three days (graphs not shown) demonstrate that the percentage of trials successful with one border-scan increased: 70%, 71%, and 74%, in the three days.

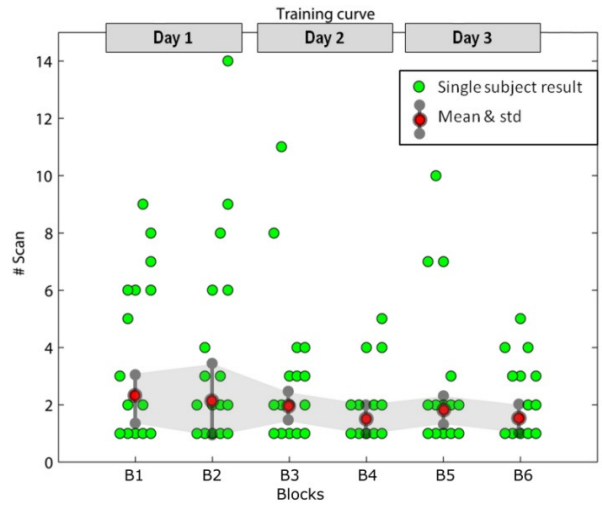



Fig. 8 Improvements in performance due to learning. The graph shows the number of scans needed to achieve correct preshaping of the hand organized as individual trials, blocks of trials, and days.

The results of the workspace tests by five subjects are presented in Table V. The Friedman test, demonstrated no statistical difference among the 6 positions ($p = 0.28$) that are at the boundaries of a 20x20x10 cm$^3$ volume, but a statistical difference among subjects ($p = 0.29$) and objects ($p < 0.001$).

TABLE V
WORKSPACE (WS) TEST: MEAN SCAN TIMES ± STD

| Object | WS centre | 10cm right | 10cm back | 10 cm left | 10 cm front | 10 cm top | Mean |
|---|---|---|---|---|---|---|---|
| **O1** | 1.4 | 1.8 | 2.4 | 1.0 | 1.7 | 1.3 | **1.6±0.0** |
| **O3** | 1.0 | 1.1 | 1.0 | 1.0 | 1.0 | 1.0 | **1.0±0.5** |
| **O5** | 2.2 | 2.3 | 1.5 | 1.8 | 2.0 | 2.5 | **2.1±0.4** |
| **O7** | 1.5 | 1.7 | 1.0 | 2.0 | 1.3 | 1.8 | **1.6±0.4** |
| **Mean** | **1.5 ±0.5** | **1.7 ±0.5** | **1.5 ±0.7** | **1.4 ±0.5** | **1.5 ±0.4** | **1.7 ±0.7** | |

## IV. DISCUSSION

The off-line data analysis and real time experimental results enlightened the potentialities and limitations of the present application, here discussed in detail.

### A. Horizontal vs. Vertical Resolution

These tests show that for subjects with normal or corrected vision and no previous training, the patterns of EOG potentials can be used to estimate horizontal gaze shifts with significant accuracy, but vertical gaze shifts with poorer or insignificant accuracy. This is true within the experimented conditions: rotation of one degree of freedom at a time (either only vertical or only horizontal eye rotation), using Equations (2) to estimate the lengths, the rotational angle range of 8-9.5 degrees and mean eye speed of 11 degrees/second.

These outcomes confirm previous findings on the asymmetry of horizontal vs. vertical smooth pursuits, which is found at all ages in humans and primates [15]-[16]. Collewijn and Tamminga attribute the horizontal-vertical asymmetry in adults to the fact that most objects that are pursued in daily life move in a horizontal plane [15], however, Grönqvist *et al.* [17] and several other studies with infants demonstrate the developmental origins of such asymmetry. Besides these possible reasons, the vertical pupil movement range is much smaller than that of the horizontal range, causing it to be much more susceptible to external perturbations such as head movements. The combination of these factors may explain the larger overshoots of the vertical scans (as shown e.g. in Fig. 5) that yielded to low or insignificant values of the Spearman's correlation coefficient.

### B. Object Recognition

The scatter plot graph and k-nn results demonstrate that the proposed 5-class object recognition problem can be solved with great success by simply applying Equations (2) to the vertical and horizontal EOG signals and thresholding the outputs (similarly to [13] and [14]). Of course, a successful recognition would also depend on how objects differ in sizes and proportions; considering two standard deviations as the distance for separating sample clouds related to different groups, we could speculate that -in the worst case- objects differing 56 mm in the vertical side and 24 mm in the horizontal

side could be successfully recognized (cf. Table III). The greatest mean errors in the vertical axis were found for those objects with shorter longer side (i.e. the 95 × 50 mm seal tape and the 38 × 38 mm ball). In such cases the errors are so large that the effective size points in the scatter plot (marked with a cross in Fig. 8), fallout from the ellipses centred on the mean estimation point (marked with circles). The reason yielding to such large systematic errors can mostly be attributed to the low vertical resolution.

### *C. On the Feasibility of the System: On-line Experiments*

The hypotheses for operation of a HAD employing the present control paradigm are: (i) no head movements while border scanning, (ii) objects shape resembling rectangles, and (iii) limited workspace (eye-target distance and eye elevation). With the on-line experiments we attempted to get insights on how such hypotheses were far from reality.

The first hypothesis fits into real life practice: while observing objects it is not uncommon that we keep our head still, and additionally, since the border scan could be completed very quickly, the head in the meantime could be considered as still (as also shown by the good results achieved in these experiments). The results from on-line experiments prove this: without constraining the head, subjects soon learnt how to operate the system to a good level. A method for eliminating head artefact in EOG recordings could be by measuring and combining eye movements with head movements.

The second hypothesis (i.e. objects resemble parallelograms or rectangles) is also realistic, since the visual representation of most objects of our life that can be held within one hand are quite small, and can be approximated to parallelograms or even rectangles. If just the short side of an object is required to correctly preshape the HAD and grasp it, the effective whole shape becomes insignificant. Additionally, it should be noted that the non-linearity of EOG signals with eye rotations greater than 20-30 degrees [7], should generate incorrect estimations. Indeed, objects that can be held in one hand must be smaller than the hand aperture size (reasonably below 15 cm), therefore will require border-scans with very limited gaze shifts (i.e. within the linear range). A further proof comes from the on-line experiment outcomes; the selected objects were objects that can be found in an everyday life scenario; they had a visual (2-d) representation different from a rectangle (e.g. the small and big bottles, the ball), however our control system demonstrated good operation. Although the number of classified outputs during the on-line experiments was limited to four preshaping postures, the proposed size-estimation algorithm based on Equations (2) is general and theoretically holds the potential to recognize more than just four grasp types.

The third hypothesis (i.e. limited workspace) is the strictest one, and cannot be assured in practice. However, since this is a trigonometric problem, it influences less the farther the distance; i.e. it generates larger errors at closer distances. In fact, the perspective causes the eye gaze shifts ($\Delta\alpha$) needed to observe an object sized $s$ to vary as the arctangent of the ratio between $s$ and the distance $d$. With larger distances, the slope of the function $\Delta\alpha=f(d)$ tends to zero, and therefore the gaze shift variations, $\Delta\alpha$ become less sensitive to distance variations. As an example, a 5 cm large wineglass distant 20 cm from the eye (which is very close to the face) yields to gaze shift variations of around 5 % every increasing/decreasing centimetre in distance. The sensitivity to distance for the same glass in a more realistic workspace, e.g. 70 cm far from eyes, decreases to about 1 % and gets even lower the farther the workspace is (but it is senseless wanting to grasp an unreachable object). As a proof of this, the results from the on-line workspace test, demonstrated no statistical differences in HAD operability among positions within a 20 × 20 × 10 $cm^3$ parallelogram, centred 71 cm far from the eyes.

The on-line control experiments showed that all fifteen subjects –although with imperfect results– finished the task when they first used the system, after just a short 10-minute training. These results enlightened the intuitiveness of the system, and demonstrated that the principle of operation can be learnt in the order of minutes, unlike other non-invasive HMI based e.g. on motor imagery (that require long training procedures). Additionally, the short-term learning experiments with five healthy subjects over three sessions showed that the performance is likely to improve with training. The present system also offers potential advantages compared to other non physiological and non-invasive control approaches: indeed it is more intimate/personal than voice-control systems [18] (a command can be issued without other people knowing), and less specific compared to HMI with a fixed number of commands, like tongue, foot or other keyboard-based control systems with a fixed number of keys [19], [20].

The objective of this study was to assess the feasibility of a new application for EOG in the assistive robotics field, potentially promising for those people with severe hand dysfunction, e.g. due to spinal cord or brachial plexus injury or degenerative neuromuscular diseases. The clinical viability of this system on the target population should be assessed on a subjective basis. It is worth underlining that all results should be assessed with regards to the simple -and practical- algorithms used. It is foreseen that more complex algorithms (still fast enough for insuring real time applicability of the system) should allow for more accurate estimation of the size of the objects, and hence finer classification.

## ACKNOWLEDGMENT

The authors would like to thank V. Monaco, J. Rigosa, R. Sassu and S.P. Murdoch for their help and the participants that volunteered in the experiments.

## REFERENCES

[1] C. Cipriani, M. Controzzi, M. C. Carrozza, "The SmartHand Transradial Prosthesis," *Journal of NeuroEngineering and Rehabilitation*, vol. 8, no. 32, 2011.

[2] L. Lucas, M. DiCicco, Y. Matsuoka, "An EMG-controlled hand exoskeleton for natural pinching," *J Robot Mechatr,* vol. 16, pp. 482-488, 2004.

[3] D. B. Popović, T. Sinkjær, M. B. Popović, "Electrical stimulation as a means for achieving recovery of function in stroke patients," *J NeuroRehabilitation,* vol. 25, no. 1, pp. 45-58, 2009.

[4] G. Pfurtscheller, C. Guger, G. Mller, G. Krausz, C. Neuper, "Brain oscillations control hand orthosis in a tetraplegics," *Neurosci Lett* Vol. 292, no. 3, pp. 211-214, 2000.
[5] L. R. Hochberg, M. D. Serruya, G. M. Friehs, J. A. Mukand, M. Saleh, A. H. Caplan, A. Branner, D. Chen, R. D. Penn, J. P. Donoghue, "Neuronal ensemble control of prosthetic devices by a human with tetraplegia," *Nature*, vol. 442, pp. 164-171, 2006.
[6] S. Micera, J. Carpaneto and S. Raspopovic, "Control of hand prostheses using peripheral information," *IEEE Reviews in Biomedical Engineering*, vol. 3, pp. 48-68, 2010.
[7] G. B. Arden, A. Barrada, J. H. Kelsey, "New clinical test of retinal function based upon the standing potential of the eye," *Br J Ophthalmol* vol. 46, pp. 449-467, 1962.
[8] J. R. LaCourse and F. C. Hludik, Jr., "An eye movement communication control system for the disabled," *IEEE Trans. Biomed. Eng.*, vol. 37, pp. 1215-1220, 1990.
[9] R. Barea, L. Boquete, M. Mazo, and E. Lopez, "System for assisted mobility using eye movements based on electrooculography," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 10, pp. 209-218, 2002.
[10] T. Eggert, "Eye movement recordings: methods," *Developments in ophthalmology*, vol. 40, p. 15-34, 2007.
[11] M. Jeannerod, "Visuomotor channels: Their integration in goal-directed prehension," *Human Movement Science* Vol. 18, pp. 201-218, 1999.
[12] M. Jeannerod, "Intersegmental coordination during reaching at natural object" In: J. Long, A. D. Baddeley, editors. Attention and performance, vol. IX. Hillsdale, New York: Erlbaum; pp. 153-69. 1981.
[13] S. Došen and D. B. Popović, "Transradial Prosthesis: Artificial Vision for Control of Prehension," Artificial Organs, vol. 35, no. 1, pp. 37-48, 2011.
[14] S. Došen, C. Cipriani, M. Kostić, M. Controzzi, M. C. Carrozza, D. B. Popović, "Cognitive Vision System for Control of Dexterous Prosthetic Hands: Experimental Evaluation," *Journal of NeuroEngineering and Rehabilitation*, vol. 7, no. 42, 2010.
[15] H. Collewijn, E. P. Tamminga, "Human smooth and saccadic eye movements during voluntary pursuit of different target motions on different backgrounds," *Journal of Physiology*, vol. 351, pp. 217-250, 1984.
[16] K. L. Grasse, S. G. Lisberger, "Analysis of a naturally occurring asymmetry in vertical smooth pursuit eye movements in a monkey," *J Neurophysiology*, vol. 67, no. 1, pp- 164-179, 1992.
[17] H. Grönqvist, G. Gredebäck, C. von Hofsten, "Developmental asymmetries between horizontal and vertical tracking," *Vision Research*, vol. 46, pp. 1754-1761, 2006.
[18] M. Su and M. Chung, "Voice-controlled human computer interface for the disabled," *Computing Contr. Eng. J.*, pp. 225–230, Oct. 2001.
[19] L. N. S. Andreasen Struijk, " An Inductive Tongue Computer Interface for Control of Computers and Assistive Devices," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 12, pp. 2594-2597, 2006.
[20] M. C. Carrozza, A. Persichetti, C. Laschi, F. Vecchi, R. Lazzarini, P. Vacalebri, P. Dario, "Biomechatronic Interface for Controlling Robots with Voluntary Foot Movements," *IEEE/ASME Trans Mech*, vol. 12, no.1, 2007.