Neurocomputing 74 (2011) 1203–1212

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

The Dorso-medial visual stream: From neural activation to sensorimotor interaction

Eris Chinellato [a,*], Beata J. Grzyb [a], Nicoletta Marzocchi [b], Annalisa Bosco [b], Patrizia Fattori [b], Angel P. del Pobil [a,c]

[a] Robotic Intelligence Lab, Universitat Jaume I, Castellón de la Plana, Spain
[b] Dipartimento di Fisiologia Umana e Generale, Università di Bologna, Italy
[c] Department of Interaction Science, Sungkyunkwan University, Seoul, South Korea

ARTICLE INFO

Available online 27 October 2010

Keywords:
Single-cell experiments
Visuomotor interaction
Eye–hand coordination
Basis function networks
Bio-inspired systems

ABSTRACT

The posterior parietal cortex of primates, and more exactly areas of the dorso-medial visual stream, are able to encode the peripersonal space of a subject in a way suitable for gathering visual information and contextually performing purposeful gazing and arm reaching movements. Such sensorimotor knowledge of the environment is not explicit, but rather emerges through the interaction of the subject with nearby objects. In this work, single-cell data regarding the activation of primate dorso-medial stream neurons during gazing and reaching movements is studied, with the purpose of discovering meaningful pattern useful for modeling purposes. The outline of a model of the mechanisms which allow humans and other primates to build dynamical representations of their peripersonal space through active interaction with nearby objects is proposed, and a detailed description of how to employ the results of the data analysis in the model is offered. The application of the model to robotic systems will allow artificial agents to improve their skills in exploring the nearby space, and will at the same time constitute a way to validate modeling assumptions.

1. Introduction

Humans and other primates possess a superior ability in dealing with objects in their surrounding space. Neuroscience research showed that areas of the primate posterior parietal cortex are able to represent the surrounding environment in a way suitable for achieving complicate coordinated movements of eyes and arms. It appears that the reference frames required to encode for visual information and for planning and monitoring the movements of different effectors are maintained contextually by population of neurons interacting through gain field effects. The way some areas of the brain perform reference frame transformations for sensorimotor interactions in the peripersonal space constitute the neuroscientific basis of this work.

The outline of a model toward the achievement of an integrated object representation is proposed in this paper, based on the active interaction of the subject with its surrounding environment. Particular importance has been given to the use of binocular data and proprioceptive information regarding eye position, critical in the transformation of sensory data into appropriate motor signals. The model is especially focused on the process of gazing and reaching toward nearby objects seen as visual targets. The second part of this work focuses on the study of the neuroscience data useful for the implementation of different visuomotor functions, and the way the insights offered by such data can be exploited in the model development. Data regarding experiments with primates on gazing and reaching movements, and referred to the posterior parietal area of the visual cortex V6A, are analyzed and discussed, with the goal of defining a detailed modeling of cortical mechanisms during the interaction of a subject with his/her environment. Hypotheses on how sensorimotor interaction abilities are achieved by employing determinate neural populations are advanced, and ways to test them with further experiments and with computational methods are proposed.

This work constitutes the first step toward a more complete attempt of providing a robot with advanced capabilities in its purposeful interaction with the environment, through active exploration and multimodal integration of the different stimuli it receives. Performing purposeful, flexible and reliable vision-based reaching toward nearby objects is a fundamental skill to pursue in order to achieve such ambitious goal. On the other hand, experiments with the robot will constitute a novel, resourceful validation ground for computational models and neuroscience hypotheses.

2. Reaching and grasping in primates

The visual cortex of the primate brain is organized in two parallel channels, called "dorsal" and "ventral" streams. The former

---

[*] Corresponding author. Tel.: +34 964 728292.
E-mail address: eris@uji.es (E. Chinellato).

0925-2312/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.neucom.2010.07.029

elaborates visual data with the main purpose of endowing the subject with the ability of interacting with his/her environment, and its tasks are often synthesized as ''vision for action''. The latter is dedicated to object recognition and conceptual processing, and thus performs ''vision for perception''. Although a tight interaction between the two streams is necessary for most everyday tasks, dorsal stream areas are more strictly related to the planning and monitoring of reaching and grasping actions [1]. In fact, dorsal visual analysis is driven by the absolute dimension and location of target objects, requiring continuous transformations from retinal data to an effector-based frame of reference. Such transformations are very likely executed by a contextual coding of the peripersonal space in different reference frames simultaneously, very likely by area V6A and its neighbors.

To give a wider prospect on the task, the brain areas more directly involved when a subject is interacting with his peripersonal space are briefly described below (refer to Fig. 1). Visual data in primates flow from the retina to the lateral geniculate nucleus (LGN) of the thalamus, and then mainly to the primary visual cortex (V1) in the occipital lobe. The two main visual pathways go from V1 and the neighbor area V2 to the posterior parietal cortex (PPC) and the inferior temporal (IT) cortex. Object information flowing through the ventral pathway passes through V3 and V4 to the lateral occipital complex (LOC), which is in charge of object recognition. The dorsal pathway can be further subdivided in two parallel streams concerned, respectively, with movement of proximal (reaching) and distal joints (grasping). The dorso-medial pathway dedicated to reaching movements includes visual area V6, visuomotor area V6A and the medial intraparietal area (MIP). The two latter areas project to the dorsal premotor cortex PMd [2]. For what concerns grasping, object related visual information flows through a dorso-lateral pathway including area V3A and the caudal intraparietal area (CIP), and then reaches the anterior intraparietal sulcus (AIP), the grasping area of the primate brain, which projects mainly to the ventral premotor area (PMv) [3]. Motor plans devised by PMd and PMv are sent to the primary motor cortex (M1) which release proper action execution signals.

The hypothesis of parallel visuomotor channels for the transport and the preshaping components of the reach-to-grasp action is well-recognized [4]. Anatomically, these two channels fall both inside the dorsal stream, and are sometimes named dorso-medial and dorso-lateral visuomotor channels [2]. Cortical area nomenclature is still controversial, and the correspondence between human and macaque studies not completely solved, but new studies confirm the duality of the reaching–grasping process [5]. According to more established nomenclature, the most important reach-related cortical areas are V6A and MIP, both receiving their main input from V6 and projecting to the dorsal premotor cortex [2,6,7]. Neural response and functional mechanisms of the dorso-lateral stream were modeled in previous works [8,9].

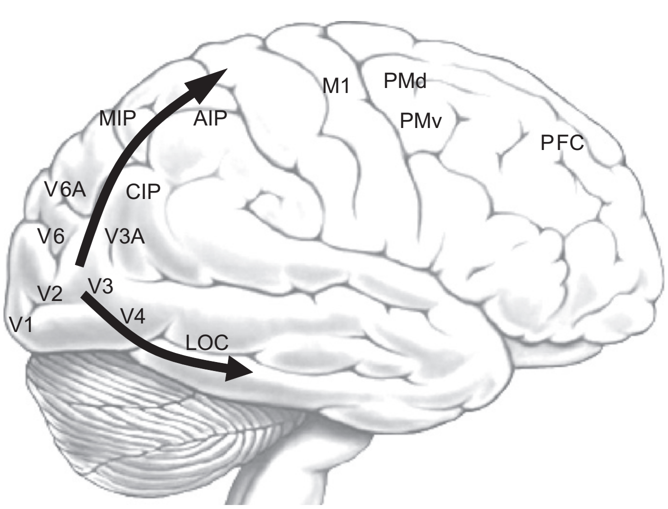

Fig. 1. The 2 visual pathways in the human brain (top arrow: dorsal; bottom arrow: ventral) with the areas involved in reaching and grasping actions.

In order to elaborate a proper action on an external target, the dorsal stream, through its two parallel sub-streams, is able of contextually manage retinal information regarding the object with proprioceptive data referred to eyes, head and hand. Area V6A very likely represents a fundamental relay station in this complex network. In fact, it employs information regarding eye position and gaze direction in order to estimate the position of surrounding objects and guide reaching movements toward them. Two types of neurons have been found in V6A which allow to sustain this hypothesis [10]. The receptive fields of neurons of the first type are organized in retinotopic coordinates, but they can encode spatial locations thanks to gaze modulation. The receptive fields of the second type of neurons are organized according to the real, absolute distribution of the subject peripersonal space. In addition, V6A contains neurons which arguably represent the target of reaching retinocentrically, and others which use a spatial representation [11]. This strongly suggests a critical role of V6A in the gradual transformation from a retinotopic to an effector-centered frame of reference. Moreover, some V6A neurons appear to be directly involved in the execution of reaching [2] and even grasping movements [12], indicating that this area is in charge of performing the visuomotor transformations required for the purposive control of arm joints, integrating visual, somatosensory and somatomotor signals in order to reach and position the hand toward a given target in the 3D space.

3. Model framework

Recent neuropsychological and neuroimaging research has shed a new light on how visuomotor coordination is organized and performed in the human brain. Thanks to such research, a model of vision-based arm movements which integrates knowledge coming from both monkey and human studies can now be advanced. A previous model we developed [13,14,8] dealt mainly with grasping issues and the planning of suitable hand configurations and contacts on target objects, leaving aside the transport component of the action. An extended framework is presented here, in which the process of reaching a visual target is thoroughly taken into account. In this work, the job of the dorsal stream is further explored. Especial interest is dedicated to subdivisions in the dorsal stream, and the way some of its areas interact, with especial interest toward area V6A, which is critical in the visual and visuomotor exploration of the environment. The proposed model outline aims at an interdisciplinary solution for providing an artificial agent with the ability of performing purposeful, flexible and reliable vision-based reaching and grasping actions on nearby objects. Such a goal is pursued while maintaining biological plausibility, and the focus on neuroscience data, for the implementation of different visuomotor functions.

3.1. Related research

Only a few neuroscientific models have been dedicated to the exploration of the mechanisms underlying the functioning of the action-related dorsal visual cortex, and the issue of integration between the contributions of dorsal and ventral pathways. The FARS model [15] attempted to emulate computationally the sensorimotor mechanisms of visual-based grasping in primates. It was implemented with neural networks, and included a large number of different brain areas (mainly inspired on monkey physiology), but only areas AIP, F5 and the primary motor cortex F1 were modeled in detail. Moreover, the FARS model was focused exclusively on grasping actions without considering the hand transport function. Herein, Jeannerod [4] proposed to extend the notion of parallel visuomotor channels about the mechanisms which operate within the dorsal stream itself, and suggested the existence of neural pathways for reaching and for grasping. These channels, although distinct, must also share a common mechanism for achieving coordination with each other. Murata and Ishida [16] also discussed parallel pathways for hand reaching and grasping. Apart from network for the control of sensory-guided hand and arm movements, their model included mechanisms of body awareness. Alternative models, such as the *multiple finger reaching* idea [17], are not given much credit, due to the quantity and quality of evidence supporting the mainstream hypothesis (see e.g. [18]).

Neuroscientific inspiration in artificial intelligence and robotics is basically confined to the use of biologically plausible artificial neural networks. Visuomotor transformations involving arm movements have been usually tackled with the use of self-organizing maps (SOM) [19,20], feedforward and recurrent neural networks [21] and less commonly with radial basis functions (RBF) [22]. In the most common approach, the system firstly learns the mapping between the image coordinates and the pan/tilt encoder coordinates of the eye motors, and then the transformation carrying from the visual input to an appropriate arm posture, suitable for reaching and grasping a target object. Some reaching models, while using biologically inspired neural networks, disregard the necessity of reference frame transformations, computing the difference vector between the target and the hand position in the eye-centered coordinate system without any additional stages [23,24].

Summarizing, although a few attempts to model the functioning of the action-related visual cortex exist, most of them do not provide any details for possible computational, and especially robotic implementation of the proposed concepts. On the other hand, biological or neuroscientific inspiration in robotics is often too superficial and conditioned by pragmatic goals and technological constraints. Aside from the theoretical contribution, the research presented here is also the first step toward the goal of improving the skills of autonomous robotic systems in their exploration of the nearby space and interaction with surrounding objects, as described below.

3.2. *A subdivision within the dorsal stream*

The whole framework of the proposed model is graphically represented in Fig. 2. After the extraction of basic visual information in V1/V2, higher level features are generated in V3 and sent to the two streams. Along the ventral stream, an increasingly invariant representation of object shape is generated in order to perform a gradual recognition of the object (areas V4 and LOC [14]). In the dorsal stream, both object shape and location have to be processed. For what concerns shape, area CIP integrates stereoptic and perspective data in order to detect pose and proportion of the target object, using also information regarding object classification [9,8]. Areas V6 and V6A estimate object location and distance, integrating retinal data with proprioceptive information about eye position. Both V3A and CIP project to AIP, which transforms object visual data in hand configurations suitable for grasping. At the same time, areas V6A and MIP determine the reaching direction and collaborate with AIP and PMd in order to execute the arm movement suitable for getting to the target object. Movement plans are devised in posterior-premotor loops, considering also the information on object identity coming from the ventral stream, and task requirements. Dorsal areas are supported by proprioceptive information coming from somatosensory areas SI/SII. The signals for action execution are sent to the motor cortex M1, and a posterior parietal-premotor-cerebellum loop is in charge of monitoring action execution in accordance to the plan.

3.3. *Toward a visuomotor knowledge of the environment*

It is often argued, and increasingly accepted by the neuroscientific community, that the dorsal stream performs all the transformations required for sensorimotor interactions exploiting gain modulation between neural populations [25]. Basis function representations, which permit to simultaneously represent stimuli in various reference frames, constitute a natural neural structure which is especially suitable for implementing gain field effects. In fact, the basis function approach has the attractive feature that both head-centric representations for arm movements and retino-centric representations for gaze movements can be encoded concurrently in the same neural map [26,27]. In this way, explicit encoding of targets in retino-centric coordinates is enhanced via

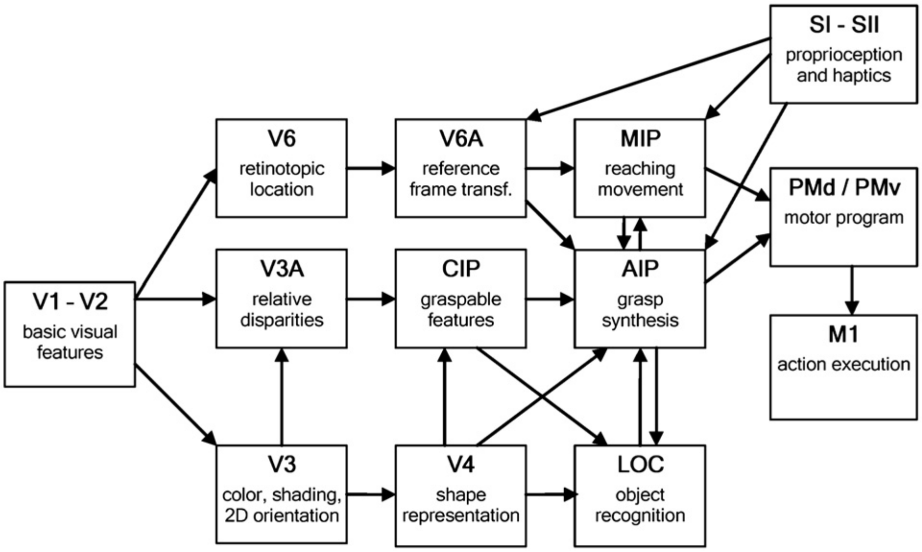

Fig. 2. Global model framework. The different information streams can be observed: the ventral stream V3-V4-LOC, the dorso-medial stream V6-V6A-MIP and the dorso-lateral stream V3A-CIP-AIP. Many more feedback connections are present, but not visualized for the sake of clarity.

gain fields to hold in parallel an implicit encoding in other reference frames [28]. Such gain fields are found in retino-centric organized eye movement areas and, most importantly, in posterior parietal area V6A [29].

Focusing on the role of V6A and the dorso-medial stream, the exploration of the environment through saccades and reaching movements constitutes the basic behavior which is employed to build an integrated visuomotor representation of the peripersonal space. Building such representation is hence done incrementally, through subsequent, increasingly complex interactions. Such learning process can be modeled by a central body-centered representation which is accessed and updated both by limb sensorimotor signals on the one hand and visual and oculomotor signals on the other hand. Indeed, this approach is consistent with findings related to area V6A, suggesting that this can be the way the dorso-medial stream solves the multiple representation problem. V6A contains neurons which code for the retinocentric position of targets, others that code for their limb-related position, and even others which seem to maintain both codings and look thus especially critical for performing sensorimotor transformations. The perception of the space and the related sensorimotor map is thus accessed and updated by visuomotor interaction, e.g. moving the gaze and the arm toward a goal position. The subject (either monkey, human, computational model or robot) performs such eye and limb movements according to either free exploration or goal directed tasks. Eyes and arms can be considered as separate effectors which receive motor control from a shared, implicit sensorimotor map of the peripersonal space, and the outcome of their contextual movements is used to update the map itself. We propose to encode the integrated sensorimotor map of the peripersonal space implicitly in a basis function structure, which models the function of V6A and its connections with purely visual and premotor areas. In the next section, neuroscience data regarding area V6A suitable for biasing the model definition and configuration are analyzed and discussed.

4. The different aspects of neural response during reaching

In previous works, single-cell experiments performed on macaque monkeys were described and analyzed [6,10,11]. This work is aimed at shedding further light on the sort of transformations performed by V6A neurons and on the coding representations they use for this purpose. The analysis approach employed here is different from the previous works, as it is performed with the final goal of achieving a computational description of V6A neurons to be used within a numerical model. In particular, the answers that need to be asked are the following. How many types of neurons does V6A contain? What are their most relevant properties and toward what tasks are they oriented? How do they perform the transformations required to coordinate and modulate retinal data, gaze direction and reaching movements?

4.1. Experiment description

The experiments analyzed here were collected at the Università di Bologna on two trained macaque monkeys. Details on the experimental protocol and related data analysis are available in a previous work [11]. They were approved by the Bioethical Committee of the University and carried out in accordance with Italian national laws and European Directives on care and use of laboratory animals. Data were collected while the monkeys were performing two possible reaching tasks toward given targets while gazing at a certain position (the fixation point).

4.1.1. Constant reaching task

In the Constant reaching protocol, the reaching target remained always in the same straight-ahead position, whereas fixation point could be in one out of three different positions (Fig. 3(a)). The monkeys performed arm movements with the contralateral limb, with the head restrained, in darkness, and maintaining steady fixation. Reaching movements transported the hand from a home-button to an LED target positioned straight-ahead (i.e. at the height of the eyes) on a fronto-parallel panel, which was located 14 cm in front of the animal. The time sequence of the reaching task is shown in Fig. 4. A trial began when the monkey decided to press the button near its chest. After 1000 ms of darkness, one of the three fixation LEDs lit up green. The monkey was required to maintain fixation on the reaching target, or on other two LED fixation points placed 3.7 cm to the right or 3.7 cm to the left of the reaching target (FIX epoch, see again Fig. 3(a)). Then, the monkey had to wait 1000–1500 ms for a change in color of the fixation LED (green to red) without performing any eye or arm movement (DELAY epoch, omitted in Fig. 4). The color change of fixation target was the go-signal for the monkey to release the home-button and perform an arm movement towards the target button, then to press it (MOV epoch). Then, the animal held its hand on the target button till fixation LED switched off, after 800–1200 ms (HOLD epoch). The switch-off of fixation LED cued the monkey to go back to the home-button to start a new trial.

Fixation of different positions was typically tested as a sequence of randomized triplets in order to collect trials in one position intermingled with the other two. Fifteen trials for each position were collected (45 trials in total).

4.1.2. Foveal reaching task

Several neurons tested with the constant reaching task underwent also another reaching task, in which the cue indicating the reaching target lit up always in the same position as the fixation target.

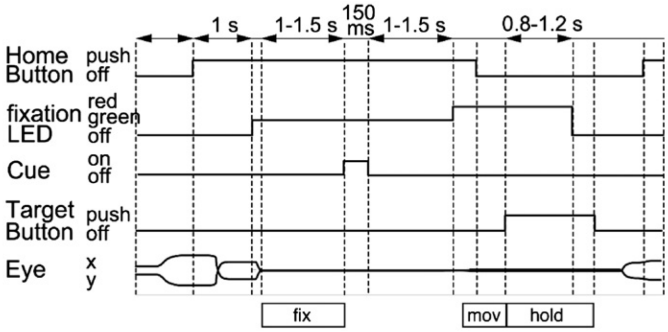

Fig. 4. Time course of a typical experimental trial of Constant or Foveal reaching.

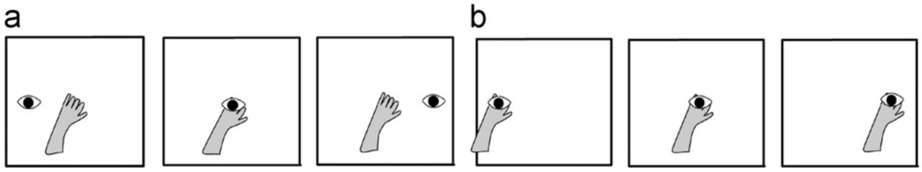

Fig. 3. Graphical description of experimental protocols. (a) Constant reaching protocol, (b) Foveal reaching protocol.

Therefore, arm reaching movements were always directed towards foveated targets (Fig. 3(b)). The retinotopic coordinates of reaching targets remained constant throughout the task, whereas the direction of movement changed trial by trial according to target position. As for the constant reaching task, foveal reaching was tested in a sequence of randomized triplets, again following the time course of Fig. 4.

Data analysis focuses on the average neural firing rate during the four time intervals of the action course (epochs FIX, DELAY, MOV and HOLD). An example of the average firing rate of one of the tested neurons during a Constant reaching experiment is plotted in Fig. 5. Neurons were classified according to their selectivity, i.e. their preferential response toward one of the three conditions for each epoch and each task. Each neuron can thus be selective in none, one or more of the four epochs; selectivity was statistically assessed by comparing the mean firing rates recorded in the three conditions (1-way ANOVA, F-test; significance level: $p<0.05$). Two types of studies were performed on the data, one on the correlation of neuronal response in different epochs and protocols, the other on a principal component analysis of their responsiveness.

4.2. Correlation analysis

As a starting point, a preferred direction index was computed for all neurons, in the two protocols and for each epoch of interest. This was done by calculating an average of the three possible positions weighted by their firing rates. The responsiveness of each neuron was thus expressed by eight values: its preferred direction in each of the four epochs of interest for both Constant and Foveal reaching protocols.

As a first insight, the responsiveness of the neural population spans the entire working range, and neuron preferred directions assume an approximately Gaussian distribution symmetrical with respect to the central direction. This can be observed in the examples of Fig. 6, referred to the Constant reaching protocol.

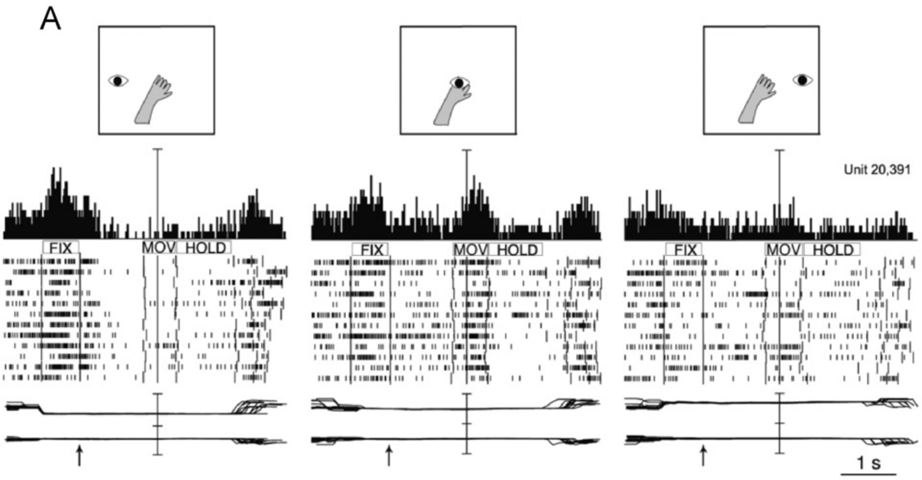

Fig. 5. Example of activation pattern of a neuron during a Constant reaching task.

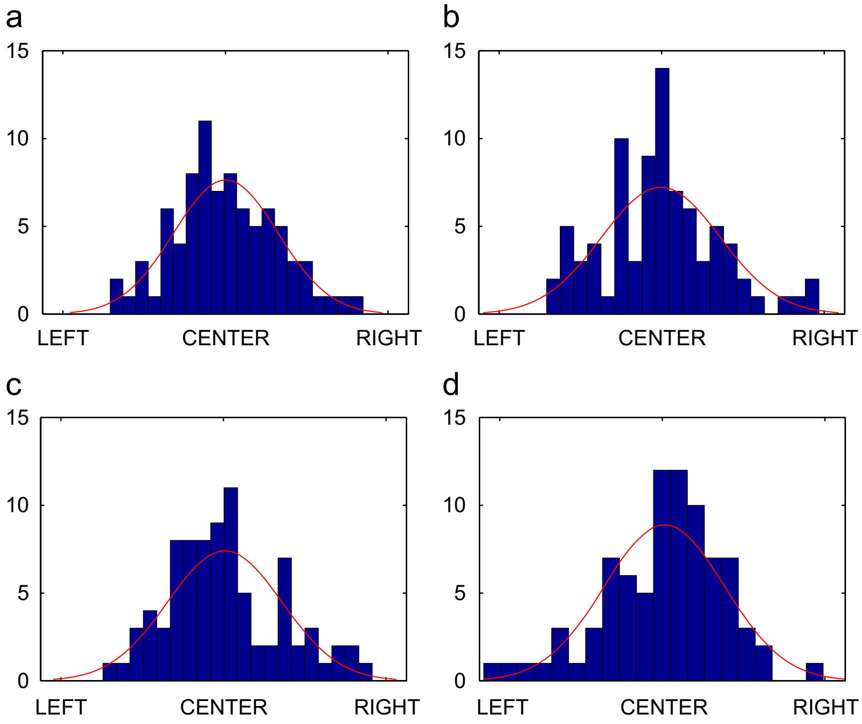

Fig. 6. Preferred direction: within epoch distributions. (a) FIX epoch, (b) DELAY epoch, (c) MOV epoch, (d) HOLD epoch.

It remains to be verified how the choice of the target positions affect such distribution, and it cannot be excluded that other neurons would be selective for positions further away from the center. In any case, the fact that all input space is covered by the neural population satisfies one of the basic requirements of basis function structures [26], and a Gaussian distribution of responsiveness around the central target indicates a homogeneous response of V6A neurons with respect to stimuli placed horizontally in the visual space.

Interesting information can be drawn from a comparative assessment of neuron preferred directions in different conditions and epochs. The results obtained comparing the preferred directions of neurons during the same epochs in the two experimental tasks are summarized in Table 1 and depicted in Fig. 7 for epochs FIX and MOV. Epoch FIX shows the highest correlation between protocols ($R = 0.69$), indicating a consistent activation across tasks during this epoch (Fig. 7(a)). On the other hand, no correlation between tasks was found for the MOV epoch ($R = -0.02$, Fig. 7(b)). Correlation coefficients for DELAY and HOLD tasks ($R = 0.36$ and $R = 0.24$, respectively), indicate some correlation, but not a strong one. These results suggest that the change in protocol affects principally the motor components of the neural responsiveness, while gaze selectivity (mainly referred to epoch FIX) is largely unaffected by the movement change. The mixed neuronal response elicited by DELAY and HOLD epochs may indicate a dual nature of this processing steps, composed of both visual and motor components. Indeed, although DELAY could appear as a gaze dominated epoch, it contains the preparation of the motion plan, and it is thus reasonable to think that an important motor component is also activated during this epoch. Similarly, the motor nature of the HOLD epoch is counterbalanced by the subject visual attention toward the Return signal, which is released while the subject holds the object. Neurons activated in both Constant and Foveal reaching during the DELAY and HOLD tasks seem to be able to encode for both gazing and reaching directions, even when they are different. Such neurons are thus good candidates for performing reference frame transformations and maintaining an integrated sensorimotor representation of the visual space. Possible correlations are more apparent if only neurons selective in one or both tasks are considered (see color-coding in Fig. 7). Epoch selectivity itself, with some neurons selective during only one epoch and other for all the course of the action, supports the above hypotheses [11].

Relevant considerations can be drawn also from the study of how neural responsiveness changes during the action course within the same experimental protocol. This can be done comparing the preferred direction of neurons in the same task but in different epochs, as in Fig. 8. The correlation coefficients obtained in this analysis are summarized in Table 2. The highest and most consistent correspondence in preferred directions can be observed between the DELAY and MOV epochs for both Constant (Fig. 8(a)) and Foveal protocols. This finding suggests a certain processing uniformity across such epochs. No correlations are found between FIX and other epochs for the Constant protocol, and only mild correlations are observed in DELAY/HOLD and MOV/HOLD (Fig. 8(b)). In Foveal reaching the situation is different, as all epochs show some correspondence. Especially the three epochs DELAY-MOV-HOLD are very well-correlated, but also FIX is now much more correlated with all other epochs (see e.g. Figs. 8(d) and (c)). Hence, when the gaze is directed where the hand is (Foveal reaching); there is a coupling in the discharge in HOLD and the epochs preceding it, and even a predominantly visual epoch, as FIX, is consistent with the reaching direction. Conversely, when the hand is maintained in a location not gazed at (Constant reaching), the cell discharge is less correlated to DELAY and MOV activity, probably because the spatial coordinates used in that stage are in a different frame of reference. Computationally, in the Foveal reaching case the target corresponds in all reference frames, whilst in

Table 1
Correlation between constant and foveal reaching for each epoch.

Epoch	Correlation coefficient
FIX	0.69
DELAY	0.36
MOV	−0.02
HOLD	0.24

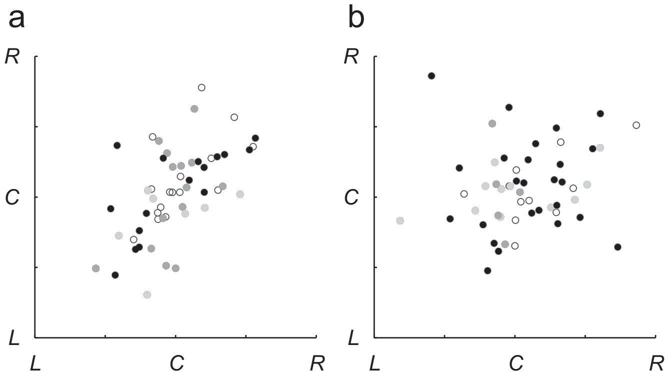

Fig. 7. Preferred direction: same epoch, different tasks (L=left; C=center; R=right). Dot color = neuron selectivity: white—not selective; light gray—selective in Constant; dark gray—selective in Foveal; black—selective in Constant and Foveal. (a) FIX epoch: Constant (x) vs. Foveal (y), (b) MOV epoch: Constant (x) vs. Foveal (y).

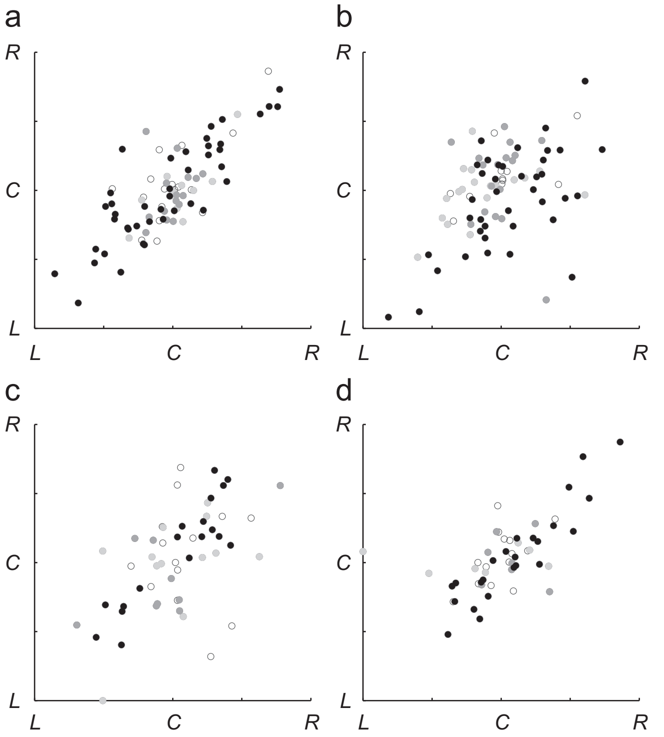

Fig. 8. Preferred direction: different epochs, same task (L=left; C=center; R=right). Dot color = neuron selectivity: white—not selective; light gray—selective in x epoch; dark gray—selective in y epoch; black—selective in both epochs. (a) Constant reaching: DELAY (x) vs. MOV (y), (b) Constant reaching: MOV (x) vs. HOLD (y), (c) Foveal reaching: FIX (x) vs. DELAY (y), (d) Foveal reaching: MOV (x) vs. HOLD (y).

Table 2
Correlation between different epochs for constant and foveal reaching.

Epochs	Constant reaching	Foveal reaching
FIX/DELAY	0.15	0.90
FIX/MOV	0.02	0.76
FIX/HOLD	0.20	0.61
DELAY/MOV	0.86	0.83
DELAY/HOLD	0.56	0.81
HOLD/MOV	0.54	0.92

Constant reaching the effectors (eye and arm) require different transformations to encode the location of their targets.

As a general pattern, some neurons seem to maintain their responsiveness across epochs and protocols, others completely change their preferred direction. These findings suggest the presence of important temporal issues, and a strong effect of action stage on neural responsiveness. A possible interpretation of this activity pattern is that some neurons sustain their activation, maybe for maintaining their coding of the target position in the head-centered reference frame. Other neurons perform instead the transformations between this frame and the retinocentric and arm-centered references when required. For modeling purposes, the first type of neurons would be in charge of maintaining the common visuomotor spatial representation, whilst the second type of neurons would be in charge of accessing and modifying such representation according to different sensorimotor events.

4.3. *Principal component analysis*

In order to better understand the sort of representation used by V6A neurons, the next step in our study was to perform a principal component analysis (PCA) on the responsiveness of all neurons and conditions (LEFT, CENTER, RIGHT) of an experimental protocol for each epoch of interest. PCA was thus executed over a 87×3 dataset for each of the four epochs. In all cases, the two first principal components accounted for nearly or more than 90% of the data variability. Thus, for both Constant and Foveal reaching, two components are almost enough to represent the whole range of the three different experimental conditions. This means that most neurons are "predictable" in their activity pattern, showing reasonably monotonic activation patterns. It would be very interesting to study those neurons which break this predictability, requiring the intervention of a third principal component, but more data are needed for this purpose. A normalized representation of the three eigenvectors obtained for each epoch during Constant and Foveal reaching is depicted in Fig. 9. The relative weights of the eigenvectors, which exemplify their capacity of representing the whole dataset, and obtained normalizing their eigenvalues, are also provided.

A first interesting aspect that can be noticed is the strict similarity between the principal components of the DELAY and MOV epochs in both protocols (Figs. 9(b) and (c)). Such finding confirms and reinforces the previously mentioned potential correlation between these two epochs. In Constant reaching, a very good correspondence can also be observed between the FIX and HOLD epochs (Figs. 9(a) and (d)), showing a relation between them which was not quite clear from the correlation graphs. For the Foveal reaching protocol (Figs. 9(e–h)), one major change is noticeable with respect to Constant reaching: while the correspondence between DELAY and MOV remains clear, epoch HOLD is now definitely closer in its principal components to these two epochs than to FIX. Indeed, correlation graphs for Foveal reaching were already showing how HOLD had a good correlation with both DELAY and MOV epochs. It is also interesting to observe how DELAY and MOV principal components remain consistent across protocols. As the reaching target remains the same in the two protocols, this result seems to confirm the dominance of arm motor control in these epochs. The correspondence between the HOLD and DELAY/MOV epochs in Foveal reaching could be explained considering that in this case the attention of the subject is directed toward the same position during DELAY (while planning the movement), MOV (while executing the movement), and HOLD (while waiting for the Return signal). In the Constant reaching task, instead, this correspondence is present for DELAY and MOV, but not for HOLD. Indeed, in the latter epoch the subject is holding its hand in one position, but its visual attention is directed toward the fixation point. From a neuroscience point of view, the similarity of the DELAY and MOV epochs in the PCA analysis might mean that the neural activity corresponding to the MOV epochs really begins during the previous epoch. This could imply that V6A neurons are strongly involved in movement planning and preparation. Still, they maintain their activation during movement execution, very likely for performing a feed-forward control loop as part of a recurrent parietal-premotor circuit, as recent anatomical studies support [30]. At the same time, this analysis suggests that a mixed basis function population with selectivity for different phases of the action across the visuomotor space can be a good approximation for modeling the job of V6A neurons.

5. Discussion

The principal component analysis of the previous section offers insights on how to define a first approximation for modeling the job of V6A neurons. Starting from the above analysis, a population of artificial neurons can be generated which is able to emulate the sort of transformations and modulations between visual data and gaze and arm movements performed by the dorso-medial stream. The different properties captured in this work can be used to tune the behavior of the neural population with various input sets corresponding to the different experimental conditions. As explained above, computational architectures based on basis functions are most suitable for modeling such behavior. A very important issue at this stage of the process is how to define the basis function neurons. Summarizing the above findings, it can be said that area V6A includes neurons having only visual response, neurons apparently involved mainly in motor actions and mixed neurons, activated in all phases of sensorimotor processes. These visuomotor neurons are very likely dedicated to allow smooth transformations from one reference frame to the other.

The model we are developing should assess what computational advantages could be obtained by such responsiveness pattern. The first modeling step is thus to learn the transformations between visual, oculomotor and arm motor space using basis function networks inspired on V6A neurons. For simplicity at this stage, only two arm joints can be used, and no tilt movements of the eyes, so that the accessible environment is a 2D space placed horizontally in front of the subject, as in most monkey experiments. The different types of V6A neurons can be simulated with populations of radial basis function neurons with receptive fields distributed according to different criteria. As a first approximation, predominantly visual neurons, such as those more strongly activated during epochs FIX and HOLD, can be distributed uniformly on the visual space; motor neurons, corresponding to epoch MOV, are better organized according to arm joint space; intermediate neurons can simulate a head-based reference frame, and can be placed according to different criteria at the same time, visual, joint and vergence/version oculomotor space. Preliminary results indicate that this neural organization is especially suitable for both direct and

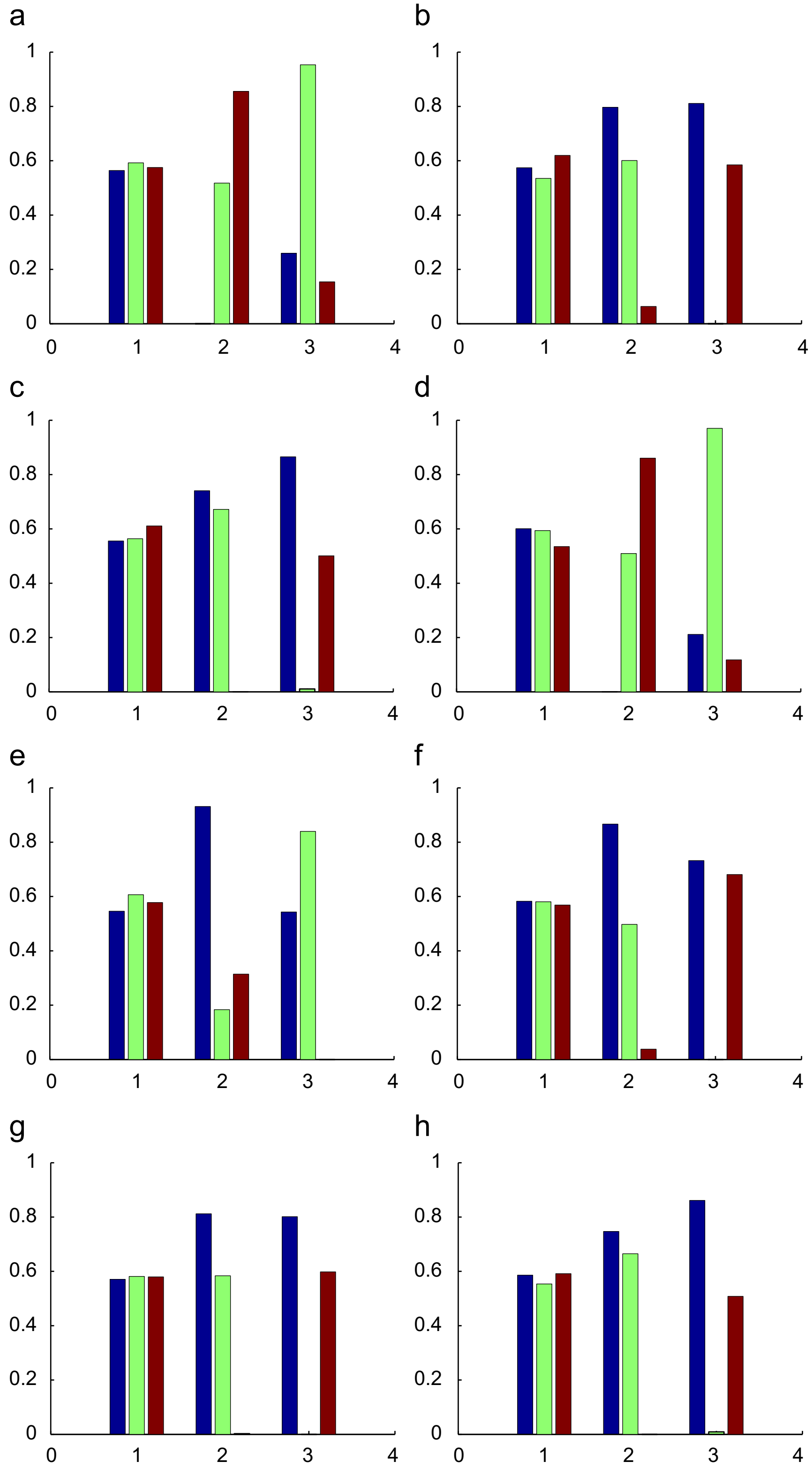

Fig. 9. PCA for Constant (above) and Foveal reaching (below). Principal components of each epoch across conditions, with correspondent weights (%). (a) FIX epoch: 80.8; 15.2; 4.0, (b) DELAY epoch: 76.0; 16.0; 8.0, (c) MOV epoch: 79.1; 13.3; 7.6, (d) HOLD epoch: 76.8; 16.4; 6.8, (e) FIX epoch: 73.1; 15.8; 11.1, (f) DELAY epoch: 73.6; 15.6; 10.8, (g) MOV epoch: 84.2; 9.5; 6.3, (h) HOLD epoch: 84.2; 10.7; 5.1.

inverse transformations between the different reference frames required for visual data gathering, eye gazing movements and arm reaching actions [31].

Data analysis and basis function modeling concord in giving V6A a critical role in the whole model framework of Fig. 2. In fact, V6A would act as a visuomotor relay station able to integrate and modulate between visual information (coming from earlier visual areas such as V2/V3), oculomotor signals (in coordination with areas LIP and FEF, not shown in Fig. 1), and arm motor signals (with other posterior parietal areas MIP and AIP and premotor areas, especially PMd). Once some more data are available, the analysis of V6A neuron activity presented in this paper could be extended using singular value decomposition (SVD), successfully applied so far to premotor data recordings [32]. Such analysis would be helpful in testing whether gaze direction and hand movement direction are separable variables for a given neuron. Computationally, the activations of the neurons from each class (separable, not separable) could be used for fitting the parameter values for different classes of neurons which use direct, intermediate or gain field encoding. Further experiments both on the neuroscience and the computational sides will help in assessing the above hypotheses, and better clarify what is the contribution of V6A in various different cortical mechanisms and the corresponding data flows.

6. Conclusions

This work described research aimed at better understanding the role of the dorso-medial visual stream in the sensorimotor transformations required for the planning and execution of gazing and reaching actions. The above analysis helps in clarifying what sort of computation is performed by dorsal stream neurons, namely those pertaining to area V6A, in order to maintain a perfect coordination between retinal data, gaze direction and arm movements. The outline of a comprehensive model, detailed for what concerns the changes in reference frames related to various sensorimotor conditions, was advanced in this paper. The model served as basis for a single-cell data analysis performed with a computational stance, which contributes to shed a new light on the role and organization of V6A neural subpopulations.

Detailed model implementation and robotic experiments will help in further clarifying the mechanisms behind eye–arm coordination and reciprocal guidance, and reference frame transformations in primates. This should carry to a better understanding of the transformations performed between retinocentric, effector-based and distance/vergence-based representations in various environments and working conditions. The predictions obtained by the model and the robotic experiments could then be tested through the development of new neuroscience studies. From a pragmatic point of view, a robot emulating the above mechanisms should be able to purposefully and consistently interact with its environment building its skills on the integration of different stimuli. Such skills would be based on the building of a plastic representation of its nearby environment, representation which can be exploited for more precise and complex interactions with the environment components. This research is thus expected to provide important advancements in both robotics and neuroscience.

Acknowledgments

Support for this research has been provided in part by the European Commission's Seventh Framework Programme FP7/2007-2013 under Grant agreement 217077 (EYESHOTS project), by the Spanish Ministerio de Ciencia y Innovación (DPI-2008-06636, FPU Grant AP2007-02565), by the Italian Ministry of University and Research, by Fundació Caixa-Castello-Bancaixa (P1-1B2008-51) and by WCU (World Class University) program through the National Research Foundation of Korea, funded by the Ministry of Education, Science and Technology (Grant no. R31-2008-000-10062-0).

References

[1] M.A. Goodale, A.D. Milner, Sight Unseen, Oxford University Press, 2004.
[2] C. Galletti, D.F. Kutz, M. Gamberini, R. Breveglieri, P. Fattori, Role of the medial parieto-occipital cortex in the control of reaching and grasping movements, Experimental Brain Research 153 (2003) 158–170.
[3] J.C. Culham, C. Cavina-Pratesi, A. Singhal, The role of parietal cortex in visuomotor control: what have we learned from neuroimaging? Neuropsychologia 44 (2006) 2668–2684.
[4] M. Jeannerod, Visuomotor channels: their integration in goal-directed prehension, Human Movement Science 18 (1999) 201–218.
[5] J.C. Culham, J.P. Gallivan, C. Cavina-Pratesi, D.J. Quinlan, fMRI investigations of reaching and ego space in human superior parieto-occipital cortex, in: R. Klatzky, M. Behrmann, B. MacWhinney (Eds.), Embodiment Ego-space and Action, Lawrence Erlbaum Associates, 2008, pp. 247–274.
[6] P. Fattori, M. Gamberini, D.F. Kutz, C. Galletti, Arm-reaching neurons in the parietal area V6A of the macaque monkey, European Journal of Neuroscience 13 (2001) 2309–2313.
[7] P. Dechent, J. Frahm, Characterization of the human visual V6 complex by functional magnetic resonance imaging, European Journal of Neuroscience 17 (2003) 2201–2211.
[8] E. Chinellato, A.P. del Pobil, Distance and orientation estimation of graspable objects in natural and artificial systems, Neurocomputing 72 (2008) 879–886.
[9] E. Chinellato, A.P. del Pobil, Neural coding in the dorsal visual stream, in: M. Asada, J. Hallam, J.-A. Meyer, J. Tanipages (Eds.), From Animals to Animats, LNAI 5040, Springer2008, pp. 230–239.
[10] P. Fattori, D.F. Kutz, R. Breveglieri, N. Marzocchi, C. Galletti, Spatial tuning of reaching activity in the medial parieto-occipital cortex (area V6A) of macaque monkey, European Journal of Neuroscience 22 (2005) 956–972.
[11] N. Marzocchi, R. Breveglieri, C. Galletti, P. Fattori, Reaching activity in parietal area V6A of macaque: eye influence on arm activity or retinocentric coding of reaching movements? European Journal of Neuroscience 27 (2008) 775–789.
[12] P. Fattori, V. Raos, R. Breveglieri, A. Bosco, N. Marzocchi, C. Galletti, The dorsomedial pathway is not just for reaching: grasping neurons in the medial parieto-occipital cortex of the macaque monkey, Journal of Neuroscience 30 (2010) 342–349.
[13] E. Chinellato, Y. Demiris, A.P. del Pobil, Studying the human visual cortex for achieving action-perception coordination with robots, in: A.P. del Pobil (Ed.), Artificial Intelligence and Soft Computing, Acta Press, Anaheim, CF, USA2006, pp. 184–189.
[14] E. Chinellato, B.J. Grzyb, A.P. del Pobil, Brain mechanisms for robotic object pose estimation, in: IEEE International Joint Conference on Neural Networks, 2008, pp. 3268–3275.
[15] A.H. Fagg, M.A. Arbib, Modeling parietal-premotor interactions in primate control of grasping, Neural Networks 11 (1998) 1277–1303.
[16] A. Murata, H. Ishida, Representation of bodily self in the multimodal parieto-premotor network, in: S. Funahashi (Ed.), Representation and Brain, Springer2007, pp. 151–176.
[17] J.B.J. Smeets, E. Brenner, M. Biegstraaten, Independent control of the digits predicts an apparent hierarchy of visuomotor channels in grasping, Behavioural Brain Research 136 (2002) 427–432.
[18] C. van de Kamp, F.T.J.M. Zaal, Prehension is really reaching and grasping, Experimental Brain Research 182 (2007) 27–34.
[19] T.M. Martinetz, H.J. Ritter, K.J. Schulten, Three-dimensional neural net for learning visuomotor coordination of a robot arm, IEEE Transactions on Neural Network 1 (1990) 131–136.
[20] M. Jones, D. Vernon, Using neural networks to learn hand-eye co-ordination, Neural Computing and Applications 2 (1994) 2–12.
[21] W. Schenck, H. Hoffmann, R. Möller, Learning internal models for eye–hand coordination in reaching and grasping, in: F. Schmalhofer, R. M. Young, G. Katz (Eds.), Proceedings of EuroCogSci 2003, 2003, pp. 289–294.
[22] M. Marjanovic, B. Scassellati, M. Williamson, Self-taught visually-guided pointing for a humanoid robot, in: International Conference on Simulation of Adaptive Behavior (SAB), 1996.
[23] G. Sun, B. Scassellati, A fast and efficient model for learning to reach, International Journal of Humanoid Robotics 2 (2005) 391–413.
[24] F. Nori, L. Natale, G. Sandini, G. Metta, Autonomous learning of 3D reaching in a humanoid robot, in: IEEE International Conference on Intelligent Robots and Systems, 2007.
[25] E. Salinas, P. Thier, Gain modulation: a major computational principle of the central nervous system, Neuron 27 (2000) 15–21.
[26] S. Pouget, A. Sejnowski, Spatial transformations in the parietal cortex using basis functions, Journal of Cognitive Neuroscience 9 (1997) 222–237.
[27] A. Pouget, L.H. Snyder, Computational approaches to sensorimotor transformations, Nature Neuroscience 3 (Suppl.) (2000) 1192–1198.
[28] S. Deneve, A. Pouget, Basis functions for object-centered representations, Neuron 37 (2003) 347–359.
[29] C. Galletti, P.P. Battaglini, P. Fattori, Eye position influence on the parieto-occipital area PO (V6) of the macaque monkey, European Journal of Neuroscience 7 (1995) 2486–2501.
[30] M. Gamberini, L. Passarelli, P. Fattori, M. Zucchelli, S. Bakola, G. Luppino, C. Galletti, Cortical connections of the visuomotor parietooccipital area V6Ad of the macaque monkey, Journal of Comparative Neurology 513 (2009) 622–642.
[31] E. Chinellato, M. Antonelli, B. Grzyb, A. del Pobil, Implicit sensorimotor mapping of the peripersonal space by gazing and reaching, IEEE Transactions on Autonomous Mental Development, submitted.
[32] B. Pesaran, M.J. Nelson, R.A. Andersen, A relative position code for saccades in dorsal premotor cortex, Journal of Neuroscience 30 (2010) 6527–6537.

Eris Chinellato received his Ph.D. in Intelligent Robotics from Jaume I University (Spain, 2008), his M.Sc. in Artificial Intelligence, with the Best Student Prize, from University of Edinburgh (UK, 2002), and his Industrial Engineering Degree from Universit'a degli Studi di Padova (Italy, 1999). His Ph.D. thesis "Visual neuroscience of robotic grasping" was among the finalists of the Georges Giralt Ph.D. Award for the Best European Ph.D. thesis in robotics in 2008. His interdisciplinary research is mainly focused, but not restricted to, the use of visual information for reaching and grasping actions in natural and artificial systems. He has published in influential journals and proceedings in robotics, neuroscience, and computational neuroscience. He has served as reviewer and program committee member for international journals, conferences, and schools, and collaborated with renowned scientists such as M. A. Goodale and R. B. Fisher.

Beata J. Grzyb received her M.Sc. Degree in Computer Science from Maria Curie–Sklodowska University in Lublin, Poland. Currently she is a Ph.D. candidate in the Robotic Intelligence Lab, Jaume I University, and has already published in several journal and proceedings. In her research, she follows the approach of cognitive developmental robotics, and tackles problems related to body representation, peripersonal space representation and perception of body effectivities, by means of synthesizing neuroscience, developmental psychology, and robotics.

Nicoletta Marzocchi received her Ph.D. in Neurophysiology (2006) and her Biomedical Engineering Degree (2002) from Universitá degli Studi di Bologna (Italy). She is author or co-author of research publications on international journals and conference proceedings in neuroscience and bioengineering. She participated in many national and European research projects, all characterized by a strong interdisciplinarity. Her research activity is aimed at understanding the functional role of the posterior parietal cortex in visuomotor behavior, mainly through the acquisition and analysis of extracellular neural signals from the posterior parietal cortex.

Annalisa Bosco received his Ph.D. in Neurophysiology from University of Bologna (Italy, 2010) and his Degree in Pharmacy from University of Bologna (Italy, 2006). Her scientific interests cover different aspects of sensorimotor integration and visuomotor behavior. She had some months experience at the INSERM of Lyon (France, 2007) and at the University of Pittsburgh (USA, 2009) where she learned analytical skills and statistical tools to analyse psychophysic and electrophysiological data. As coauthor, she has published in influential journals and attended some international meetings (FENS 2008, Switzerland; SFN 2009, USA).

Patrizia Fattori obtained her Ph.D. in Neuroscience from University of L'Aquila in 1994. She was appointed Assistant Professor at the Faculty of Pharmacy of the University of Bologna in 1993; there, in 2001, she became Associate Professor. She has more than 15 years of experience with single-cell recording in awake monkeys. She has about 50 publications in high impact factor neuroscientific journals. Her research focuses on monkey studies aimed at clarifying the neural and cognitive processes that stand behind the visuomotor coordination between visual information, gaze direction and arm movements. She has collaborated in interdisciplinary research with robotics labs. She is interested in the neural mechanisms that lead from vision to action, specifically when prehension actions are involved.

Angel P. del Pobil is Professor of Computer Science and Artificial Intelligence at Jaume I University (Spain), and founder director of the Robotic Intelligence Laboratory. He holds a B.S. in Physics (Electronics, 1986) and a Ph.D. in Engineering (Robotics, 1991), both from the University of Navarra (Spain). His Ph.D. Thesis was the winner of the 1992 National Award of the Spanish Royal Academy of Doctors. Prof. del Pobil is author or co-author of over 120 research publications, and has been invited speaker of 34 tutorials, plenary talks, and seminars. His past and present research interests include: motion planning, visually-guided grasping, service robotics, mobile manipulators, visual servoing, learning for sensor-based manipulation, and the interplay between neurobiology and robotics.