Eye-Hand Coordination for Reaching in Dorsal Stream Area V6A: Computational Lessons

Eris Chinellato[1], Beata J. Grzyb[1], Nicoletta Marzocchi[2], Annalisa Bosco[2],
Patrizia Fattori[2] and Angel P. del Pobil[1]

[1] Robotic Intelligence Lab
Universitat Jaume I, Castellón de la Plana, Spain
{eris,grzyb,pobil}@uji.es
[2] Dipartimento di Fisiologia Umana e Generale
Universita' di Bologna, Italy
{nicoletta.marzocchi,annalisa.bosco,patrizia.fattori}@unibo.it

Abstract. Data related to the coordination and modulation between visual information, gaze direction and arm reaching movements in primates are analyzed from a computational point of view. The goal of the analysis is to construct a model of the mechanisms that allow humans and other primates to build dynamical representations of their peripersonal space through active interaction with nearby objects. The application of the model to robotic systems will allow artificial agents to improve their skills in their exploration of the nearby space.

1 Introduction

Despite the growing interest of robotics researchers in biologically-inspired approaches, robot vision-based reaching and grasping systems usually work on a very different level of abstraction if compared with plausible computational models of the corresponding neural mechanisms.

A previous model we developed [1, 2] 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. This paper is a part of an extended framework in which the process of reaching toward a visual target is thoroughly taken into account. The research presented here 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.

The focus of this work is on the study of the neuroscience data useful for the implementation of different visuomotor functions. Data regarding experiments with primates on gazing and reaching movements, and referred to the dorsal stream area V6A, are analyzed and discussed, with the goal of defining a detailed modeling of dorsal stream mechanisms during the interaction of a subject with his/her environment. The conclusions of such analysis are useful for both robotic applications and neuroscience research.

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 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 [3]. 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.

The correct coupling between the reaching and grasping movements, often neglected in robotic applications, is instead a fundamental and largely studied aspect in human grasping, and various plausible models on the relation between reaching and preshaping movements have been developed [4]. The hypothesis of parallel visuomotor channels for the transport and the preshaping components of the reach-to-grasp action is well recognized [5]. Anatomically, these two channels fall both inside the dorsal stream, and are sometimes named dorso-medial and dorso-lateral visuomotor channels [6]. Cortical areas 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 [7]. 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[6, 8, 9]. For what concerns the dorso-lateral stream and the control of distal joints, the caudal intraparietal sulcus CIP is dedicated to the extraction and description of visual features suitable for grasping purposes. Its neurons are strongly selective for the orientation and proportion of visual stimuli, represented in a viewer-centered way, and they were modeled in a previous work [10]. Action plans are very likely devised by the anterior intraparietal sulcus AIP, the grasping area of the primate cortex, in collaboration with premotor areas.

In order to elaborate a proper action on an external target, the dorsal stream requires two main inputs, the object shape and pose and its location with respect to the eyes and thus to the hand. These inputs are obtained by integrating retinal information regarding the object with proprioceptive data referred to eyes, head and hand. All this information is managed contextually by the dorsal stream, through its two parallel sub-streams, dorso-medial and dorso-lateral. Area V6A seems to represent a fundamental relay station in this complex network. The assumption is that information regarding eye position and gaze direction is employed by V6A in order to estimate the position of surrounding objects and guide reaching movements toward them. Two types of neurons have been found in V6A that allow to sustain this hypothesis [11]. 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 that arguably represent the target of reaching retinocentrically, and others that use a spatial representation [12]. 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 movements [6], indicating that this area is in charge (probably together with MIP) of performing the visuomotor transformations required for the purposive control of proximal arm joints, integrating visual, somatosensory and somatomotor signals in order to reach a given target in the 3D space.

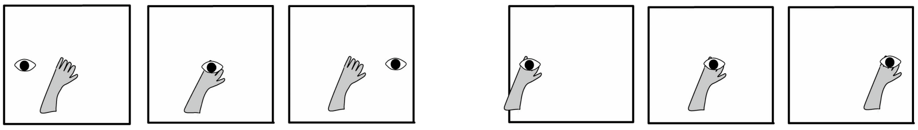
3 The Different Aspects of Neural Response During Reaching

In previous works, single-cell experiments performed on macaque monkeys were described and analyzed [8, 11, 12]. In this work we aim at shedding further light on the sort of transformations performed by V6A neurons and on the coding representations they use to 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 model that will be applied to a real robotic setup. 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?

The main repercussion of assuming this different analysis approach is that more quantitative, global measures will be favored upon classification and labeling solutions. Neurons will still be classified according to their selectivity, but their responsiveness will be quantitatively measured and compared. As a consequence, statistical analysis will be reduced and simplified, and results will be observed from a more empirical and application-oriented point of view. For example, statistical tests will not be discussed in this paper, as they were largely performed in the previous works, and some interesting conclusion will be drawn directly from visual inspection of charts and graphs.

3.1 Experiment Description

The experiments analyzed here were collected at the Università di Bologna on two trained macaque monkeys. 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 given targets while gazing at a certain position (the fixation point) illuminated by an LED (Figure 1). In the first task (**Constant reaching**) the target remained always in the same straight-ahead position, whereas the fixation point could be in one out of three different positions, as symbolized in Figure 1(a). In the second task (**Foveal reaching**) the fixation point changed in one out of three positions as in the first task, but arm-reaching movements were always directed towards the fixation point, as depicted in Figure 1(b). For other details regarding experimental procedures see [12].



(a) Constant reaching protocol

(b) Foveal reaching protocol

Fig. 1. Graphical description of experimental protocols.

The data analysis focuses on the average neural firing rate during four time intervals of the action course (epochs). The time epochs taken into account were defined as follows:

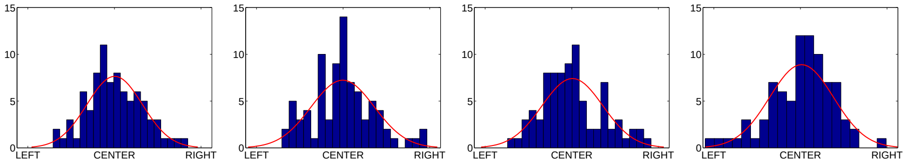
- FIX: steady fixation of the LED; starts when the gazing on the fixation point is detected and ends at the onset of the position cue indicating the position to be reached;
- DELAY: delay period before the go-signal; starts 300ms after the position cue offset and ends at the go-signal.
- MOV: arm reaching movement; starts 200ms before movement onset and lasts until movement end.
- HOLD: object holding period; starts at movement end and finishes 200 ms before return movement onset.

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 main types of analysis were performed on the data, one based on the preferred response of neurons, the other on a principal components analysis of their responsiveness.

3.2 Preferred Direction

The first step of this analysis was to compute for all neurons a preferred direction index, 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 8 values: its preferred direction in each of the 4 epochs of interest for both Constant and Foveal reaching protocols.

Figure 2 shows histograms of the responsiveness of all analyzed neurons during the 4 epochs of interest, for Constant reaching experiments. Very similar results, not plotted for space reasons, were obtained for the Foveal reaching protocol. From the results exemplified in Figure 2 it looks reasonable to assume that the responsiveness of the neural population spans the entire working range, and that neurons preferred directions assume an approximately Gaussian distribution symmetrical with respect to the central direction. 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. As neurons were sampled from both hemispheres, we checked for possible laterality effects performing the above analysis in an ipsilateral/contralateral representation instead that in a LEFT/RIGHT one. Activation histograms were completely symmetrical, confirming that no significant laterality effects could be observed, and for this reason we continued our study only considering the LEFT/RIGHT representation.

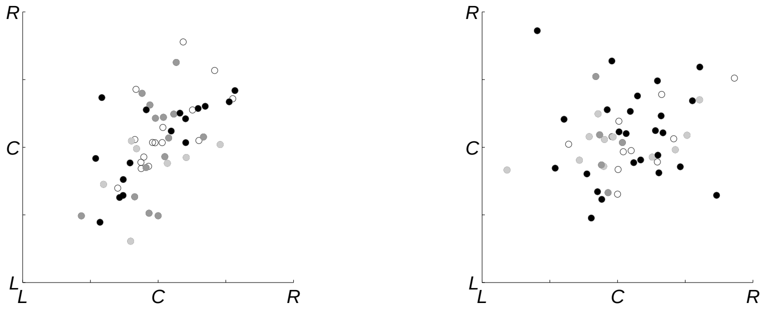


(a) FIX epoch (b) DELAY epoch (c) MOV epoch (d) HOLD epoch

Fig. 2. Preferred direction: within epoch distributions.

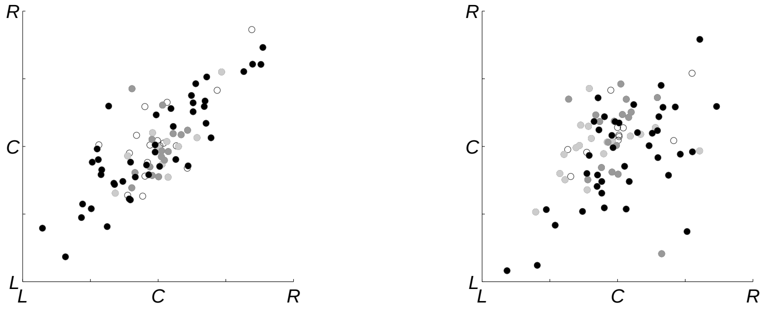
More interesting insights can be drawn from a comparative assessment of neurons preferred directions in different conditions and epochs. The results obtained comparing the preferred directions of neurons during the same epochs in the two experimental tasks are depicted in Figure 3. It can be observed how neural activation during the FIX epoch (Figure 3(a)) is rather consistent across tasks. For what concerns the MOV epoch (Figure 3(b)), instead, it is hardly possible to detect any clear correlation among tasks. 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. DELAY and HOLD epochs elicit mixed neuronal response (not shown), maybe indicating a dual nature, composed of both visual and motor components. Possible correlations are more apparent if only neurons selective in one or both tasks are considered (see color-coding in Figure 3).

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 a strong motor components is 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.

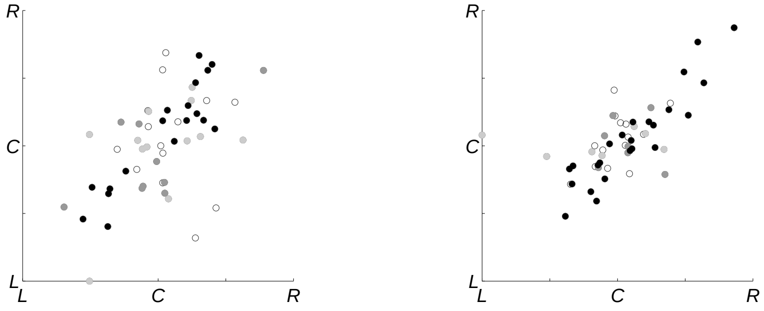


(a) FIX epoch: Constant (x) vs. Foveal (y) (b) MOV epoch: Constant (x) vs. Foveal (y)

Fig. 3. 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) 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)

Fig. 4. 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.

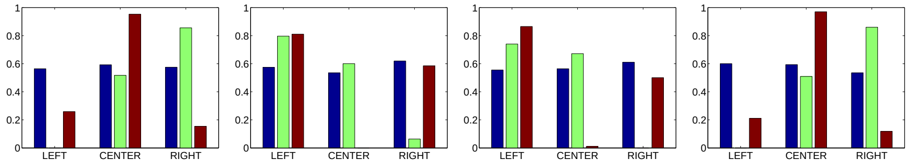
Relevant considerations can be drawn also by the study of how neural responsiveness changes during the action course within the same experimental protocol. This can be done comparing the preferred direction of neurons in the same task but in different epochs, as in Figure 4. The most apparent correspondence in preferred directions can probably be observed between the DELAY and MOV epochs for both Constant (Figure 4(a)) and Foveal protocols, suggesting a certain processing uniformity across such epochs. No other clear correlations can be observed for the Constant protocol, and the situation resembles Figure 4(b). In Foveal reaching the situation is different, as all epochs show some correspondence, and especially the three epochs DELAY-MOV-HOLD are quite well correlated, as can be seen for example in Figure 4(d) and to a minor extent in Figure 4(c). This could indicate than, 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. Conversely, when the hand is maintained in a location not gazed at (Constant reaching), the cell discharge can be uncorrelated to DELAY and MOV activity probably because the spatial coordinates used in that stage are in a different frame of reference.

In general, some neurons seem to maintain their responsiveness across epochs and protocols, others completely change their preferred direction. These findings suggests 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, whereas others perform transformations according to the mutual situation of target, eyes and hand, and action stage.

3.3 Principal Components 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 components analysis of 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 87x3 dataset for each epoch, and 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 that break this predictability, requiring the intervention of a third principal component, but more data are needed to this purpose. A normalized representation of the three eigenvectors obtained for each epoch during Constant and Foveal reaching is depicted in Figure 5. 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 the MOV epochs (Figures 5(b)

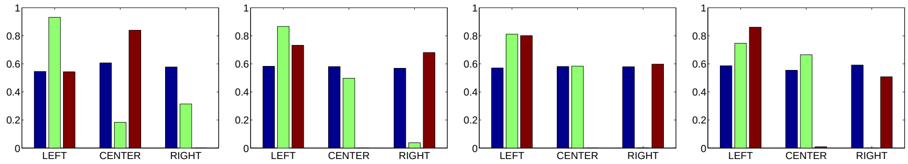


(a) FIX epoch: 80.8; 15.2; 4.0

(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

Fig. 5. PCA for Constant (above) and Foveal reaching (below). Principal components of each epoch across conditions, with correspondent weights (%).

and 5(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 (Figures 5(a) and 5(d)), showing a relation between them that was not quite clear from the correlation graphs. For the Foveal reaching protocol (Figures 5(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. The correspondence between the HOLD and DELAY/MOV epochs in the Foveal task and not in the Constant reaching task could be explained considering that in the first 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 second 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.

A different PCA analysis, performed for the 87 neurons across 4 epochs for each experimental condition, reinforces the idea that epochs can indeed be split in two groups only, and still explain most data variability. In fact, the 2 principal components of such analysis always accounted for 90% or more of the data variation. Given the results of the first PCA analysis, depicted in Figure 5, it seems reasonable to assume that a major reduction is obtained thanks to the similarity of the DELAY and MOV epochs and the FIX and HOLD epochs in Constant reaching, and to the group DELAY-MOV-HOLD in Foveal reaching.

From a neuroscience point of view, this 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 [13].

The principal components obtained in this analysis constitute a first approximation for modeling the job of V6A neurons. Starting from such components, a population of artificial neurons can be generated which is able to emulate the sort of transformation and modulation between visual data and gaze and arm movements performed by the dorso-medial stream. The different properties captured in this work will be used to tune the behavior of the neural population with various input sets corresponding to the different experimental conditions. Candidate computational architectures for modeling such behavior are Radial Basis Functions, which emulate gain field mechanisms [14], and dynamical Self Organizing Maps [15], especially suitable to the unsupervised learning of different concurrent stimuli patterns. In either case, the computational structure should be able to endow a robot with the capacity of learning the characteristics of its nearby space through active exploration.

4 Summary and Conclusions

This work described research aimed at a better understanding of the role of the dorso-medial visual stream in the planning and execution of 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. This research is expected to provide important advancements in both robotics and neuroscience.

A robot emulating dorsal stream 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 that can be exploited for more precise and complex interactions with the environment components.

Robotic experiments would help in further clarifying the mechanisms behind eye-arm coordination and reciprocal guidance and reference frame transformations in primates. A first interesting test is to extend the one-dimensional nature of the experiments presented in this work first to 2D and then, to the full 3D space, adding depth information processing and check how the mutual modulation between retinal data, gaze direction and reaching movements is required to change to adapt to the different cases. 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.

Acknowledgments

Support for this research has been provided in part by the European Community's Seventh Framework Programme FP7/2007-2013 under grant agreement 217077 (EYESHOTS project), by Ministerio de Ciencia y Innovación (FPU grant AP2007-02565) and by Fundació Caixa-Castello (project P1-1B2005-28).

References

1. E. Chinellato. *Visual Neuroscience of Robotic Grasping.* PhD thesis, Universitat Jaume I, Spain, 2008.
2. E. Chinellato, Y. Demiris, and A. P. del Pobil. Studying the human visual cortex for achieving action-perception coordination with robots. In A.P. del Pobil, editor, *Artificial Intelligence and Soft Computing*, pages 184–189. Acta Press, Anaheim, CF, USA, 2006.
3. M. A. Goodale and A. D. Milner. *Sight Unseen.* Oxford University Press, 2004.
4. R. Shadmehr and S. P. Wise. *The computational neurobiology of reaching and pointing: A foundation for motor learning.* MIT Press, 2005.
5. M. Jeannerod. Visuomotor channels: Their integration in goal-directed prehension. *Human Movement Science*, 18(2):201–218, June 1999.
6. C. Galletti, D. F. Kutz, M. Gamberini, R. Breveglieri, and P. Fattori. Role of the medial parieto-occipital cortex in the control of reaching and grasping movements. *Experimental Brain Research*, 153(2):158–170, November 2003.
7. J. C. Culham, J. P. Gallivan, C. Cavina-Pratesi, and D. J. Quinlan. fMRI investigations of reaching and ego space in human superior parieto-occipital cortex. In M. Behrmann & B. MacWhinney R. Klatzky, editor, *Embodiment, Ego-space and Action*, pages 247–274. Lawrence Erlbaum Associates, 2008.
8. P. Fattori, M. Gamberini, D. F. Kutz, and C. Galletti. 'Arm-reaching' neurons in the parietal area V6A of the macaque monkey. *Eur J Neurosci*, 13(12):2309–2313, Jun 2001.
9. P. Dechent and J. Frahm. Characterization of the human visual V6 complex by functional magnetic resonance imaging. *Eur J Neurosci*, 17(10):2201–2211, 2003.
10. E. Chinellato and A. P. del Pobil. Neural coding in the dorsal visual stream. In *International Conference on the Simulation of Adaptive Behavior*, 2008.
11. P. Fattori, D. F. Kutz, R. Breveglieri, N. Marzocchi, and C. Galletti. Spatial tuning of reaching activity in the medial parieto-occipital cortex (area v6a) of macaque monkey. *Eur J Neurosci*, 22(4):956–972, Aug 2005.
12. N. Marzocchi, R. Breveglieri, C. Galletti, and P. Fattori. Reaching activity in parietal area V6A of macaque: eye influence on arm activity or retinocentric coding of reaching movements? *Eur J Neurosci*, 27(3):775–789, Feb 2008.
13. M. Gamberini, L. Passarelli, P. Fattori, M. Zucchelli, S. Bakola, G. Luppino, and C. Galletti. Cortical connections of the visuomotor parietooccipital area V6Ad of the macaque monkey. *J Comp Neurol*, 513(6):622–642, Feb 2009.
14. S. Deneve and A. Pouget. Basis functions for object-centered representations. *Neuron*, 37(2):347–359, January 2003.
15. A. Rauber, D. Merkl, and M. Dittenbach. The growing hierarchical self-organizing map: exploratory analysis of high-dimensional data. *IEEE Transactions on Neural Networks*, 13(6):1331–1341, Nov. 2002.