Cogn Process (2012) 13 (Suppl 1):S141–S146
DOI 10.1007/s10339-012-0478-4

SHORT REPORT

Space coding for sensorimotor transformations can emerge through unsupervised learning

Michele De Filippo De Grazia · Simone Cutini · Matteo Lisi · Marco Zorzi

Abstract The posterior parietal cortex (PPC) is fundamental for sensorimotor transformations because it combines multiple sensory inputs and posture signals into different spatial reference frames that drive motor programming. Here, we present a computational model mimicking the sensorimotor transformations occurring in the PPC. A recurrent neural network with one layer of hidden neurons (restricted Boltzmann machine) learned a stochastic generative model of the sensory data without supervision. After the unsupervised learning phase, the activity of the hidden neurons was used to compute a motor program (a population code on a bidimensional map) through a simple linear projection and delta rule learning. The average motor error, calculated as the difference between the expected and the computed output, was less than 3°. Importantly, analyses of the hidden neurons revealed gain-modulated visual receptive fields, thereby showing that space coding for sensorimotor transformations similar to that observed in the PPC can emerge through unsupervised learning. These results suggest that gain modulation is an efficient coding strategy to integrate visual and postural information toward the generation of motor commands.

Keywords Neural network · Generative model · Sensorimotor transformations · Gain modulation · Parietal cortex

---

M. De Filippo De Grazia · S. Cutini (✉) · M. Lisi · M. Zorzi
Department of General Psychology, Center for Cognitive Science, University of Padova, Padua, Italy
e-mail: simone.cutini@unipd.it

Introduction

Programming goal-directed motor behavior requires a series of computational steps, ranging from the sensory acquisition of the target to the generation of the motor command, including a transformation between different coordinate reference frames. Such transformation is mandatory because the reference frames of the sensory input usually differ from those of motor effectors. The neural substrates of sensorimotor transformations are undisputedly attributed to the posterior parietal cortex (PPC) (Colby and Goldberg 1999). The PPC has been the subject of extensive research since a seminal neurophysiological investigation (Mountcastle et al. 1975) describing how its neural activity, characterized by both sensory and motor properties, is related to eye and limb movements; crucially, the PPC has been described as a sensorimotor interface for the generation of visually guided movements (Buneo and Andersen 2006). PPC neurons are functionally segregated for different types of movement (Fig. 1). In particular, the intraparietal sulcus (IPS) represents the cross-road where different sensory inputs converge to encode-specific motor programs (Buneo and Andersen 2006). The IPS contains several distinct sub-regions, each one devoted to a specific task: for instance, the lateral intraparietal area (LIP) is specialized for saccadic eye movements, the medial intraparietal area (MIP) is specialized for reaching, and the anterior intraparietal area (AIP) is involved in grasping (Sakata and Taira 1994).

Spatial representations for motor programming are distributed in a group of PPC neurons whose activity approximates a multiplicative combination of visual and posture signals (Andersen et al. 1985). Such interplay between information coming from different modalities, usually termed gain modulation, can be generally described

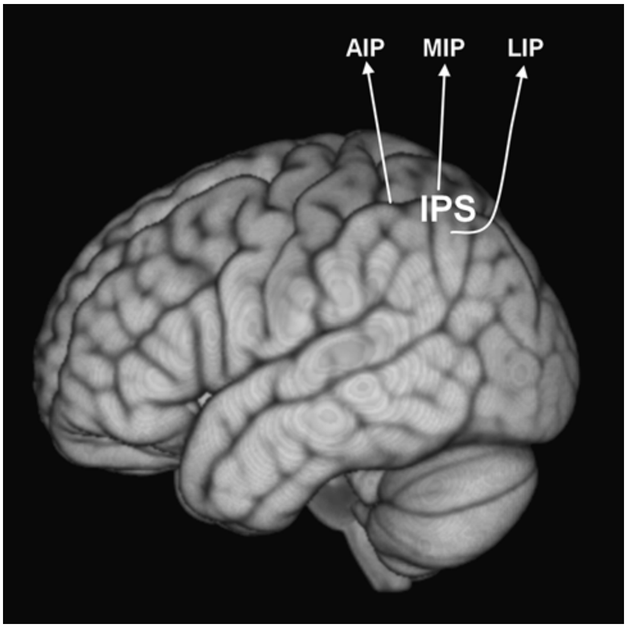


Fig. 1 The PPC of the human brain. In particular, three regions of the intraparietal sulcus (*IPS*) are crucially involved in sensorimotor transformations: anterior (*AIP*), medial (*MIP*) and lateral (*LIP*) intraparietal areas

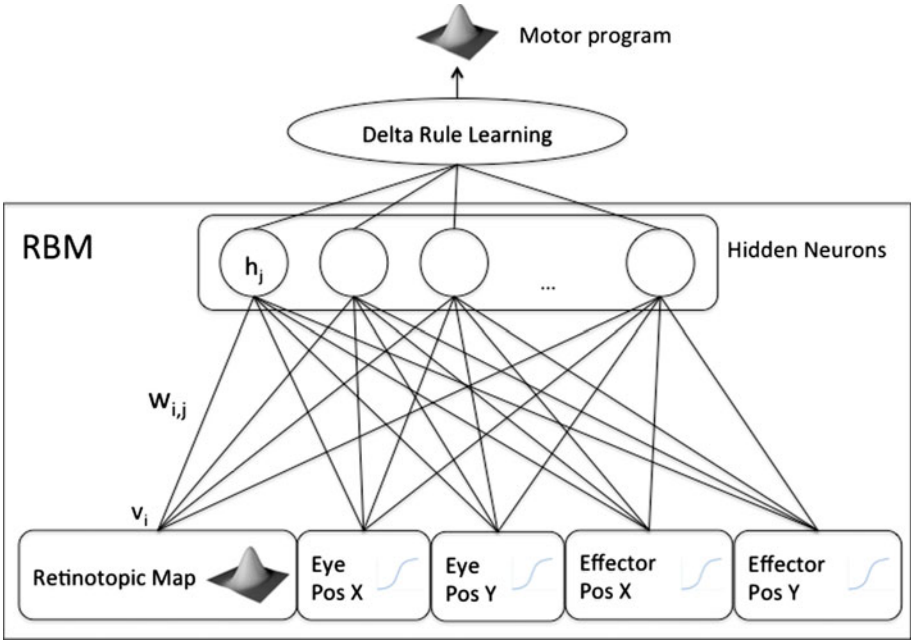
as a change in response amplitude of a neuron independently of its selectivity or receptive field characteristics. Computational investigations allowed to confirm the crucial role of gain modulation in sensorimotor transformations. In a seminal work (Zipser and Andersen 1988), a multi-layer neural network was trained with the back-propagation algorithm to transform retinal signal into head-centered coordinates. Notably, the hidden neurons of the network developed gain-fields similar to those described in parietal cortex. In another computational investigation (Mazzoni and Andersen 1991), a neural network was trained to represent visual space in head-centered coordinates with a reinforcement learning rule, which is more biologically plausible: also in this case, the authors found gain-fields in the hidden neurons. Thus, these studies show that gain modulation might provide an efficient solution to the coordinate transformation problem.

Here, we investigated whether space coding based on the gain field mechanism may emerge when learning is only concerned with efficient coding of the sensory information (i.e., unsupervised learning of a generative model) rather than finding a function that maps sensory information onto motor programs (i.e., supervised learning). We used a restricted Boltzmann machine (RBM) (Hinton et al. 2006), a stochastic recurrent neural network that learns a generative model of the sensory input without supervision and develops distributed nonlinear representations at the level of the hidden neurons during training.

Method

The model is based on a RBM, which is formed by one layer of visible (sensory) neurons and one layer of hidden neurons (feature detectors) that are fully connected by bidirectional symmetric weights (Fig. 2). The visible neurons simulated

Fig. 2 Model architecture. The RBM was used to learn without supervision a distributed nonlinear representation of the sensory input in the hidden layer. Delta rule learning was then used to compute the motor program from the activity of the hidden neurons



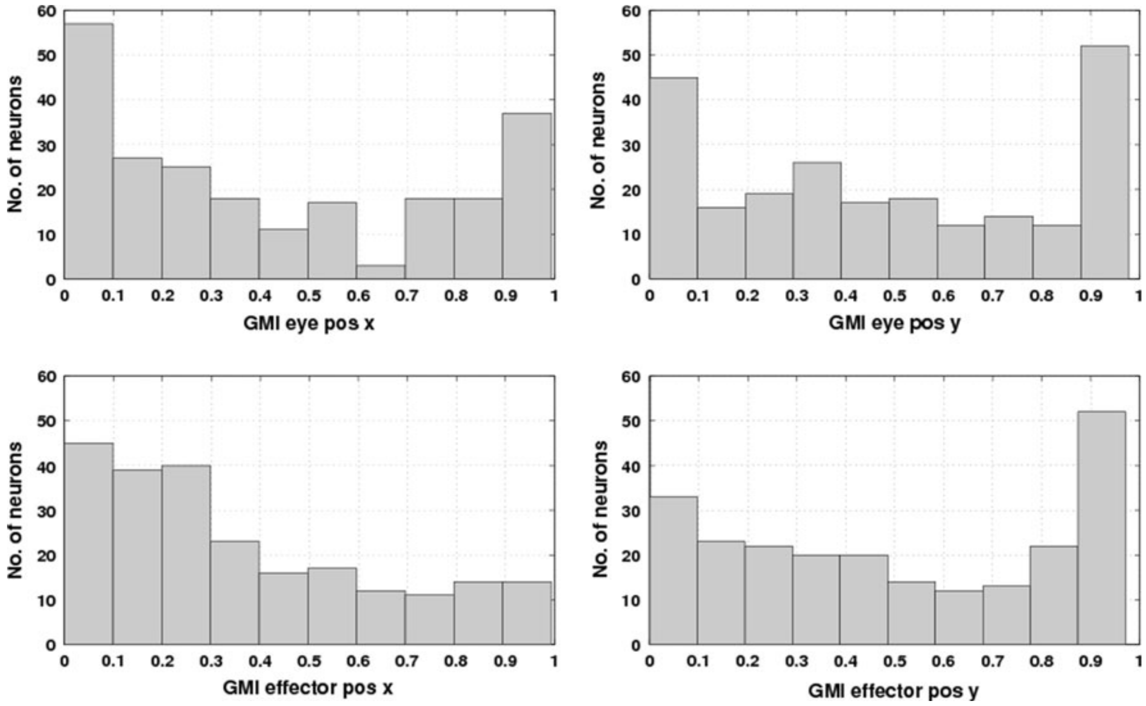
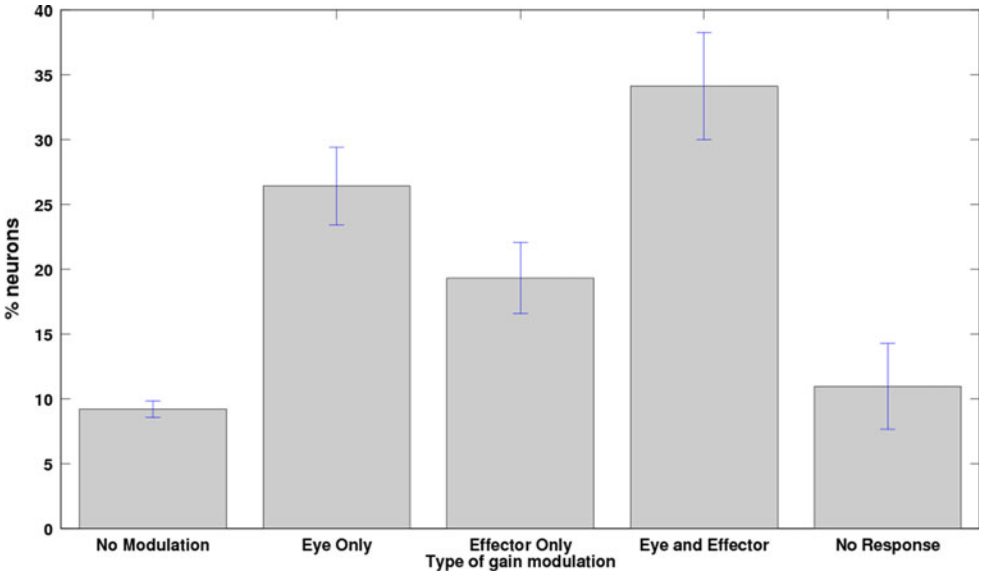


Fig. 3 Distribution of gain modulation index (GMI) values of hidden neurons in one exemplar network as a function of postural position

Fig. 4 Percentage of the hidden neurons showing a specific type of postural modulation



the activity of the cortical areas supplying sensory information to the PPC (see below), while hidden neurons were supposed to develop response properties similar to those of PPC neurons during learning. The RBM was trained with the Contrastive–Divergence learning algorithm (Hinton and Salakhutdinov 2006) to learn a generative model of the input data without supervision (i.e., maximizing the likelihood of reconstructing the data). In the "positive" phase, the visible neurons are clamped to an input vector v_i^+, and their activity spreads to the feature detectors h_j^+. In the "negative" phase, a stochastically selected binary state of the feature detectors (using their state h_j^+ as probability to turn them on) feeds back to the visible neurons through the top-down weights (i.e., reconstruction of the input vector) and then feeds forward again to the feature detectors h_j^-. The weights $w_{i,j}$ are updated with a small learning fraction ε of the difference between pairwise correlations measured in the positive and negative phases:

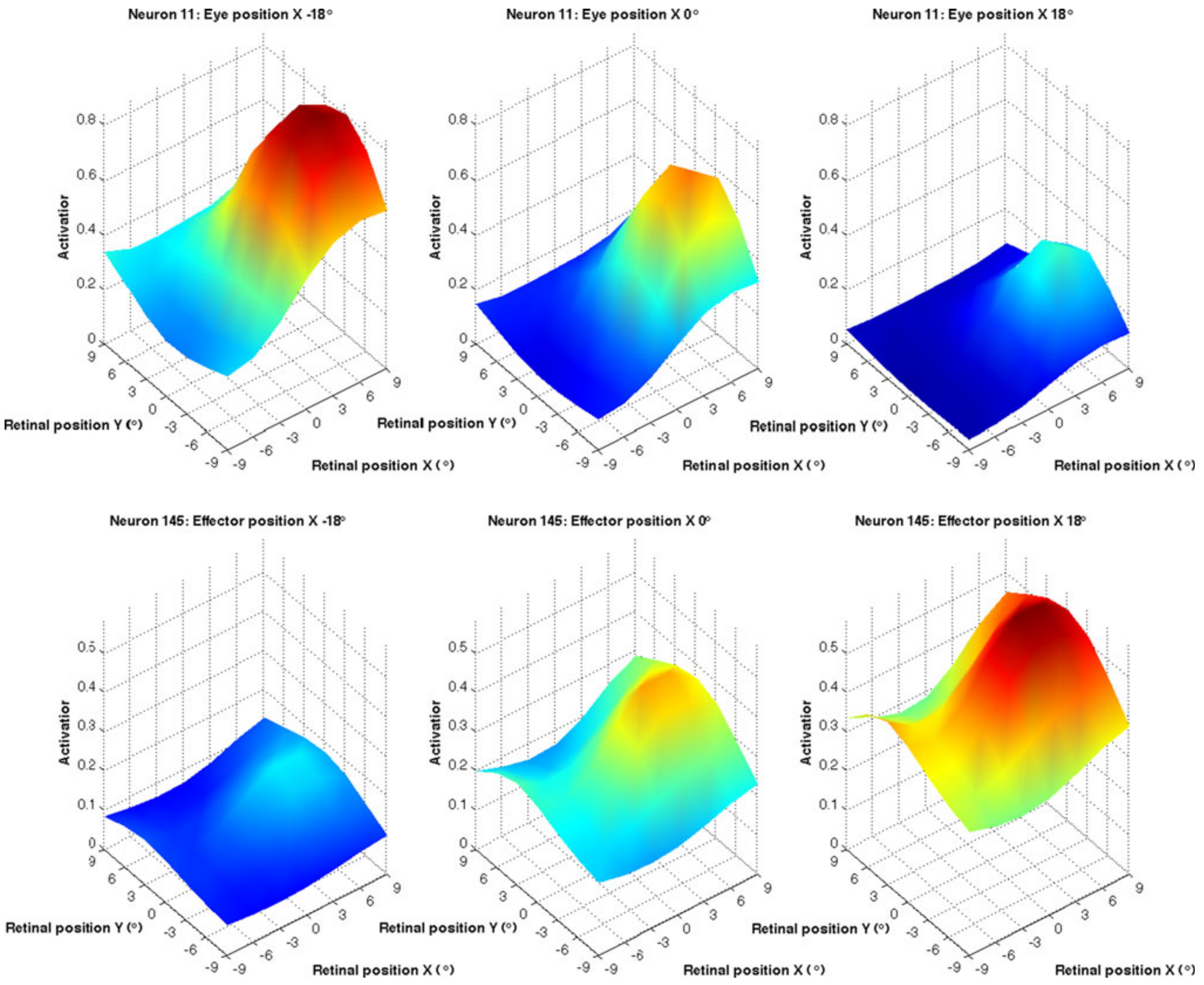


Fig. 5 Analysis of hidden layer neurons: *top row*, receptive field of Neuron 11 (Network 1) as a function of eye position and *bottom row*, receptive field of Neuron 145 (Network 1) as a function of effector position

The visible layer was composed of 357 neurons:

1. A retinotopic map, simulating area V6 (Galletti et al. 2001), consisting in a square matrix of 17 × 17 neurons with Gaussian tuning functions. Visual receptive fields were uniformly spread between −9° and +9° in increments of 3°, both in the horizontal and vertical dimensions;
2. Four postural maps (each one consisting in 17 neurons) encoding the horizontal and vertical position of the eye and the horizontal and vertical position of an effector (e.g., hand). The neurons used a sigmoid activation function to encode position between −18° and 18°, in steps of 3°.

This input representation is broadly consistent with neurophysiological data, and the basic characteristics of the neural network are in line with those used in earlier computational investigations of sensorimotor transformations (Pouget and Snyder 2000). The hidden layer was composed of 250 neurons, and its activity reflected a distributed nonlinear encoding of the sensory input. The training set presented for each training epoch consisted in 520 patterns balanced for each visual and postural position. In the learning phase, the training set was presented for a total of 3,000 epochs.

After training the RBM, we assessed whether the activity of hidden neurons supported the computation of motor programs through a simple linear projection (Pouget and Snyder 2000). The motor program (for each training pattern) was a population code on a map of 35 × 35 neurons that coded the target position in coordinates centered on the effector. Thus, the RBM's hidden layer was the input to the motor map (Fig. 2). Delta rule learning, which minimizes the error through gradient descent, was

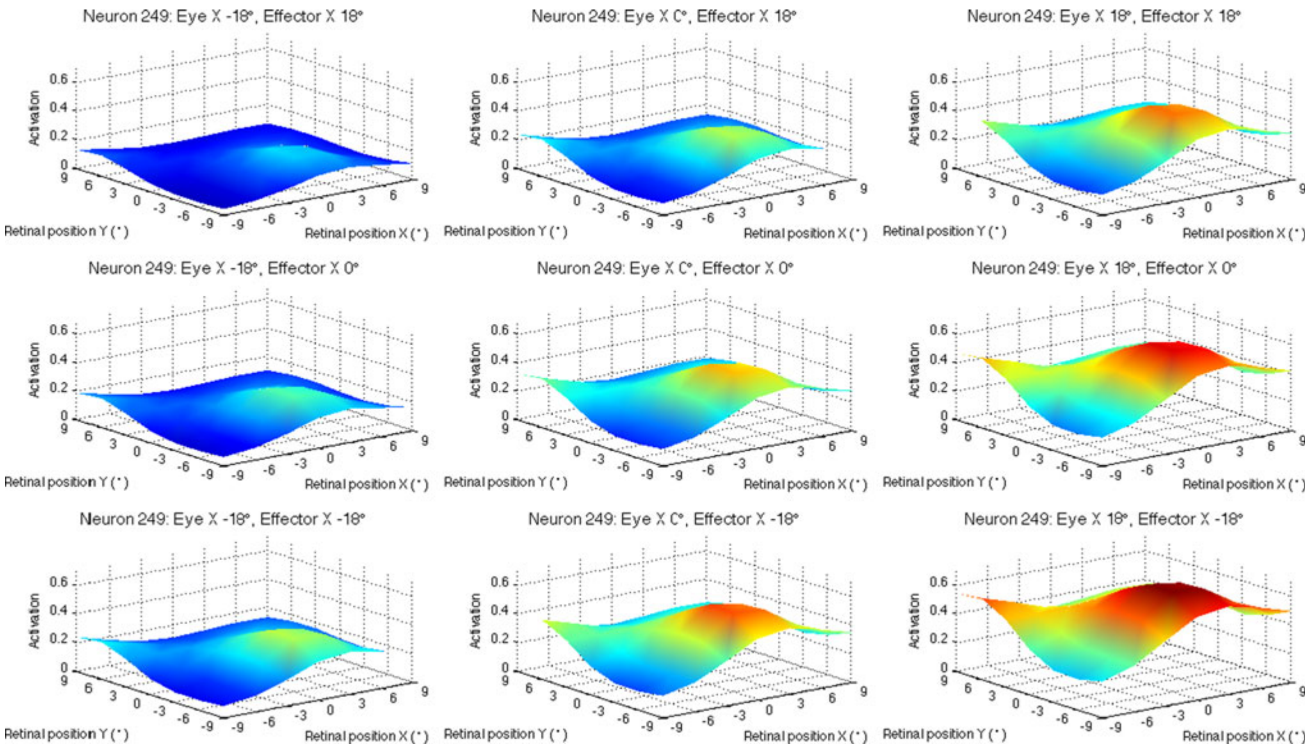


Fig. 6 Analysis of hidden neurons: a more complex receptive field of Neuron 249 (Network 1) modulated by both eye and effector position

used to adjust the connection weights between hidden neurons and motor neurons. The model's performance was tested by computing the average motor error, consisting in the difference between expected and actual output. In order to achieve reliable results, we trained ten networks with different initial random weights.

Activity in the hidden layer was analyzed with the approach that is usually adopted in neurophysiological studies (e.g., Andersen et al. 1985) to detect the presence of gain modulation in parietal neurons: the modulation of the visual receptive field of each hidden neuron was mapped by recording its response to each target location as a function of eye and effector positions (Pouget and Snyder 2000). For each postural position, the normalized ratio between maximum and minimum volume profiles (changing a postural position at a time) was computed. The postural positions that generated the maximum volume profile have been set as constant values for the invariant postural positions. Thus, for each hidden neuron, four values indicating the gain modulation index (GMI) were obtained.

Results

The average motor error (calculated as the mean error of the ten trained networks) was below 4°, the typical performance error threshold adopted while training monkeys in neurophysiological investigations of intraparietal neurons (Zipser and Andersen 1988).

The investigation of the hidden layer revealed a small number of neurons (about 10 %) characterized by close-to-nil activations. Most of the remaining hidden neurons exhibited clear gain modulation. The distribution of GMI values across hidden neurons is shown in Fig. 3: a different GMI distribution can be observed for each postural variable.

The GMI scores are distributed along the range [0,1], where zero means that the receptive field of a neuron is unaffected by postural information, hence varying postural positions does not generate any modulation, while high GMI scores indicate that the receptive field of a neuron is strongly modulated by a postural variable. As illustrated in the figure, all postural variables exhibited GMI values encompassing the entire range. Then, we classified all hidden neurons of the ten trained networks in terms of gain modulation effect in each postural position. We used a GMI of 0.5 as an arbitrary threshold to calculate the percentages of neurons falling in different categories according to modulation type (see Fig. 4).

Only about 9 % of the neurons responded to the visual stimuli but were not modulated by postural information. Indeed, activity of most neurons was modulated by the position of the eye (26 %), the effector (19 %), or both

(34 %). We investigated the neurons' receptive fields to verify the gain modulation effect and to understand which reference frame was used to compute the motor program. Notably, all receptive fields of the hidden neurons turned out to be strictly retinotopic. The top row of Fig. 5 shows the receptive field of Neuron 11 in the hidden layer of Network 1 as a function of different horizontal eye positions (−18°, 0° and 18°). Although the neuron's receptive field is fixed on the retina, its activity is modulated by eye position: its largest response is observed when the eye is positioned at −18°. The bottom row of Fig. 5 shows the receptive fields of Neuron 145 of Network 1 as a function of different horizontal effector positions: here, the largest response is observed when the effector is positioned at 18°.

Finally, Fig. 6 shows the receptive field activity of Neuron 249 of Network 1 as a function of different horizontal eye and effector positions. These examples show that there are complex selectivity patterns for different combinations of eye and effector positions. For this neuron, the largest response is observed for horizontal eye position of 18° and horizontal effector position of −18°.

Conclusions

Our results show that spatial representations for visually guided movements can emerge through unsupervised learning in a generative model that predicts the sensory input via top-down activation. Generative models are extremely appealing because they represent plausible models of cortical learning and are consistent with neurobiological theories that emphasize the mixing of bottom-up and top-down interactions in the brain (Hinton 2007). Indeed, generative models can account for empirical data at both behavioral and neurophysiological level (Stoianov and Zorzi 2012). Our study sheds light on the integrative mechanisms involved in sensorimotor transformations. The interaction between visual and postural signals observed in the present computational model is in line with neurophysiological data recorded in PPC (Salinas and Sejnowski 2001). All the receptive fields of the hidden neurons were fixed on the retina, showing that their reference frame is based on retinotopic coordinates; nevertheless, their activity was markedly modulated by postural variables: both these properties are strikingly similar to the neural properties observed in the sub-regions of PPC (Brotchie et al. 1995). Most importantly, these properties spontaneously emerged in the hidden layer even though learning did not involve any coordinate transformations (cf. Zipser and Andersen 1988; Mazzoni and Andersen 1991).

Our results show that gain modulation is an efficient coding strategy to integrate visual and postural information, independently from the generation of motor commands. Importantly, gain-fields have been also described in cortical and subcortical areas that are not specifically involved in coordinate transformation (Salinas and Sejnowski 2001), thereby suggesting that gain modulation can be considered a more general brain mechanism supporting a broader class of nonlinear transformations. In conclusion, the present study shows that gain modulation spontaneously emerges as a mechanism for multisensory integration in a generative model that simply learns to efficiently encode the sensory data without supervision.

Acknowledgments This study was supported by a grant # 210922 from the European Research Council to M.Z.

Conflict of Interest This supplement was not sponsored by outside commercial interests. It was funded entirely by ECONA, Via dei Marsi, 78, 00185 Roma, Italy.

References

Andersen RA, Essick GK, Siegel RM (1985) The encoding of spatial location by posterior parietal neurons. Science 230:456–458

Brotchie P, Andersen R, Snyder L (1995) Head position signals used by parietal neurons to encode locations of visual stimuli. Nature 375:232–235

Buneo CA, Andersen RA (2006) The posterior parietal cortex: sensorimotor interface for the planning and online control of visually guided movements. Neuropsychologia 44(13):2594–2606

Colby CL, Goldberg ME (1999) Space and attention in parietal cortex. Ann Rev Neurosci 22:319–349

Galletti C, Gamberini M, Kutz DF, Fattori P, Luppino G, Matelli M (2001) The cortical connections of area V6: an occipito-parietal network processing visual information. Eur J Neurosci 13(8): 1572–1588

Hinton GE (2007) Learning multiple layers of representation. Trends Cogn Sci 11(10):428–434

Hinton GE, Salakhutdinov RR (2006) Reducing the dimensionality of data with neural networks. Science 313(5786):504–507

Hinton GE, Osindero S, Teh YW (2006) A fast learning algorithm for deep belief nets. Neural Comput 18(7):1527–1554

Mazzoni P, Andersen RA (1991) A more biologically plausible learning rule for neural networks. Neurobiology 88(May): 4433–4437

Mountcastle VB, Lynch JC, Georgopoulos A, Sakata H, Acuna C (1975) Posterior parietal association cortex of the monkey: command functions for operations within extrapersonal space. J Neurophysiol 38(4):871–908

Pouget A, Snyder LH (2000) Computational approaches to sensorimotor transformations. Nat Neurosci 3 Suppl(november):1192–1198

Sakata H, Taira M (1994) Parietal control of hand action. Curr Opin Neurobiol 4(6):847–856

Salinas E, Sejnowski TJ (2001) Gain modulation in the central nervous system: where behavior, neurophysiology, and computation meet. Neuroscientist 7(5):430–440

Stoianov I, Zorzi M (2012) Emergence of a "visual number sense" in hierarchical generative models. Nat Neurosci 15(2):194–196

Zipser D, Andersen RA (1988) A back-propagation programmed network that simulates response properties of a subset of posterior parietal neurons. Nature 331(6158):679–684