

Decoding the Cortical Transformations for Visually Guided Reaching in 3D Space

Gunnar Blohm¹⁻⁴, Gerald P. Keith^{1,5} and
J. Douglas Crawford^{1,5,7}

¹Centre for Vision Research, York University, Toronto, Ontario M5P 2L3 Canada, ²Centre for Systems Engineering and Applied Mechanics and Laboratory of Neurophysiology, Université Catholique de Louvain, 1348, Louvain-la-Neuve, Belgium, ³Centre for Neuroscience Studies, Faculty of Arts and Science and ⁴Department of Physiology, Queen's University, Kingston, Ontario, K7L 3N6, Canada, ⁵Department of Psychology, York University, Toronto, Ontario, M5P 2L3, Canada, ⁶Departments of Biology and ⁷Departments of Kinesiology and Health Sciences, York University, Toronto, Ontario, M5P 2L3, Canada

To explore the possible cortical mechanisms underlying the 3-dimensional (3D) visuomotor transformation for reaching, we trained a 4-layer feed-forward artificial neural network to compute a reach vector (output) from the visual positions of both the hand and target viewed from different eye and head orientations (inputs). The emergent properties of the intermediate layers reflected several known neurophysiological findings, for example, gain field-like modulations and position-dependent shifting of receptive fields (RFs). We performed a reference frame analysis for each individual network unit, simulating standard electrophysiological experiments, that is, RF mapping (unit input), motor field mapping, and microstimulation effects (unit outputs). At the level of individual units (in both intermediate layers), the 3 different electrophysiological approaches identified different reference frames, demonstrating that these techniques reveal different neuronal properties and suggesting that a comparison across these techniques is required to understand the neural code of physiological networks. This analysis showed fixed input-output relationships within each layer and, more importantly, within each unit. These local reference frame transformation modules provide the basic elements for the global transformation; their parallel contributions are combined in a gain field-like fashion at the population level to implement both the linear and nonlinear elements of the 3D visuomotor transformation.

Keywords: eye movements, head movements, arm movements, pointing, model, neural network, visuomotor transformation

Introduction

Reaching toward an object in 3-dimensional (3D) space requires a transformation of visual signals into a motor plan suitable to drive the arm (Flash and Sejnowski 2001; Blohm and Crawford 2007). At the heart of this process is a “reference frame transformation” that converts eye-centered sensory signals (also often called gaze-centered or retinotopic signals) into shoulder-centered motor signals (Soechting et al. 1991; Snyder 2000; Crawford et al. 2004; Blohm and Crawford 2007). When reference frame transformations are analyzed from a 2D perspective that utilizes mathematics appropriate for translations, they can be trivialized as a sequence of vectorial movement commands that are independent of the initial, intermediate, or final frame of references (Jurgens et al. 1981; Goldberg and Bruce 1990; Crawford and Guitton 1997; Buneo et al. 2002). However, 3D geometry includes both translational and rotational aspects that require complex and nonlinear solutions (Pouget and Sejnowski 1997; Blohm and Crawford

2007). For example, to compute an accurate reach plan from visual signals, the brain needs to account for 3D eye and head orientation, the spherical geometry of the eyes, as well as for the offset between the centers of rotation of eyes, head, and shoulder (Crawford et al. 2000; Henriques and Crawford 2002; Blohm and Crawford 2007). These 3D computations are not a side issue that can be “tacked onto” a 2D stream; they are the central problem in sensorimotor reference frame transformations (Crawford et al. 2004).

Such transformations are not merely of theoretical interest; they pose a practical problem that needs to be solved for proper behavior. Failure to account for eye and head orientation would lead to reach errors—potentially quite large—whenever the eyes and head are not pointed straight ahead in an upright orientation. For example, if the head is tilted torsionally (Fig. 1A) or if gaze is simply deviated in an oblique direction (Fig. 1B), failure to account for the resulting distortions of retinal projection and their complex relation to shoulder orientation will lead to errors in both reach direction and depth (Blohm and Crawford 2007). Because such large errors are not observed behaviorally (Soechting et al. 1991; Henriques et al. 1998, 2003; Henriques and Crawford 2002; Medendorp and Crawford 2002; Blohm and Crawford 2007), the brain must take into account the full complexity of the body geometry.

At the moment, no one knows how the brain implements these transformations for 3D reach. A number of theoretical studies have investigated the visuomotor transformation using 1D or 2D approximations (Zipser and Andersen 1988; Salinas and Abbott 1995, 1996, 2001; Pouget and Snyder 2000; Xing and Andersen 2000; Deneve et al. 2001; Mascaro et al. 2003; Smith and Crawford 2005), but as discussed above, these approximations do not capture the complexity of the real transformation. Similarly, numerous electrophysiological experiments have investigated the visuomotor transformations for reach from a 2D perspective (for reviews, see Snyder 2000; Battaglia-Mayer et al. 2003). These experiments have provided critical insights into the reach-related neural signals in parietal and frontal cortex. However, without a proper 3D theoretical framework, one cannot have a complete understanding of the existing data or design optimal experiments.

For example, many physiological and theoretical investigations of reference frame transformations have focused on the analysis of “gain fields,” that is, the eye/head position-dependent modulation of visual and motor receptive field (RF) amplitudes (e.g., Andersen et al. 1985; Brotchie et al. 2003). Theoretical gain fields were first observed in artificial

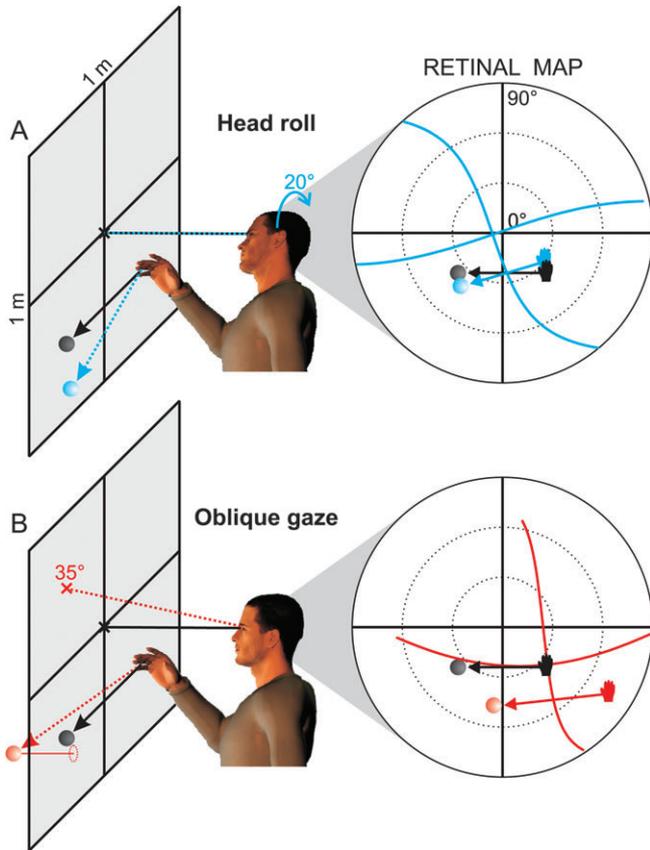


Figure 1. Nonlinearity of the 3D reference frame transformation. The retinal map (right panel) shows the projection of the hand and target (left panel, black dot) as well as the screen horizontal and vertical onto the retina. (A) Head roll (20° clockwise, blue) with eyes straight ahead. The black lines show the reference position with eye and head straight ahead. (B) Oblique gaze (35° up-left on oblique axis) with head straight ahead. The screen horizontal/vertical lines on the retinal projection (right side) show the nonlinearities arising for secondary eye and/or head position. The colored arrows and dots on the left panel show the movement that would have been generated from the retinal projection of the real hand and target position (left panel, black) if eye/head positions had been ignored. (Adapted from Blohm and Crawford [2007].)

neural nets trained to transform a 2D location on an eye-centered map into a 2D location on a head- or space-centered map (Zipser and Andersen 1988). However, this transformation bears little resemblance to the geometric transformations required for 3D reach (Blohm and Crawford 2007). Moreover, when reference frame transformations are reduced to 2D (i.e., linear, additive, and commutative) processes, gain fields are not a theoretical necessity (Pouget and Sejnowski 1997). This has led to the suggestion that gain fields are not related to reference frame transformations but rather serve some other function (Colby and Goldberg 1999).

Further, there is reason to suspect that the computations required for 3D geometry—spherical projections, nonlinear noncommutative multiplicative transformations, misaligned centers of rotation—will necessitate entirely different implementations in real or artificial networks than for 2D computations. With the addition of 3D constraints, one cannot assume that properties that arose from 2D simulations will hold up, and neither can one assume that the arguments against them will hold. This is a question that is best answered empirically.

Another important question is whether the intermediate layers of neural networks involved in sensorimotor transformations use any coherent reference frame at all. Networks that were designed to use basis function units have shown convincingly that a 2D reference frame transformation (e.g., from eye coordinates to head coordinates) can be done using intermediate units that employ mixed, intermediate frames (Pouget and Sejnowski 1997; Pouget and Snyder 2000; Xing and Andersen 2000). After all, it is only the output of the network that matters for behavior, not the intermediate stages. However, it has not been shown if the same network behavior arises in nonbasis function networks that are trained to perform a 3D transformation.

Finally, a question of critical importance to experimentalists relates to the nature of the information that can be derived using standard electrophysiological methods: microstimulation and the correlation of neuronal activity to either sensory or motor parameters. For example, often there is an implicit assumption that visual RFs, motor tuning, and stimulation-evoked movement should align in an optimal visuomotor transformation. Misalignments are often treated as “noise” or technical limitations. However, several theoretical studies have provided results that question these basic assumptions (Pellionisz and Llinas 1985; Zipser and Andersen 1988; Pellionisz and Ramos 1993; Smith and Crawford 2001, 2005). As we will demonstrate below, there is good reason to suspect that the units within a network involved in a 3D reference frame transformation must simultaneously encode different types of information (related to both sensory input and motor output) in different reference frames and that different electrophysiological techniques reveal different aspects of these codes.

A complete model of the sensorimotor transformations for reach would include multisensory representations of both target and hand position (Sober and Sabes 2003, 2005; Ren et al. 2006) and a complete model of limb dynamics (Todorov 2000; Todorov and Jordan 2002) including feedback control loops at different levels. However, our main goal here was to model the early feed-forward parietal-frontal transformations from visual inputs into motor commands, with a focus on the role of extraretinal eye and head position signals. Therefore, we have restricted our representations of both target and hand position inputs to visual coordinates and our outputs to motor commands in shoulder coordinates. We believe that this is experimentally justifiable because 1) visual representations appear to override proprioceptive representations of hand position (Sober and Sabes 2003, 2005), 2) there is evidence that target and hand position signals are compared in visual coordinates in parietal cortex (Buneo et al. 2002), and 3) parietal cortex is not thought to be involved in the detailed control of limb dynamics (Kalaska and Crammond 1992). Thus, here we are simply asking: how do neural networks transform visual inputs into the early motor plan for 3D reach?

We recently modeled this transformation using explicit geometric transformations (Blohm and Crawford 2007), but “black box” models cannot show how neural networks solve the problem. Given that the real transformation appears to occur accurately in a feed-forward fashion (Blohm and Crawford 2007), it is reasonable to develop the necessary framework using a feed-forward artificial neural net. A similar approach was used with some success with the 3D visuomotor transformation for saccades (Smith and Crawford 2005), but the transformations for reach are much more complex. To date,

no one has trained an artificial net to solve the 3D geometry required for accurate reaching.

Figure 2 provides an overview of the approach that we took in the current study. We began (Fig. 2A) with the black box model of the 3D transformations for reach that we developed in our previous study (Blohm and Crawford 2007). Briefly, a visual desired movement vector has to be rotated and translated into a shoulder-centered motor command. Rotations have to account for eye-in-head and head-on-shoulder orientation, whereas translations account for the fact that the centers of rotation of the eyes do not coincide with that of the head and the center of rotation of the head does not coincide with that of the shoulder. We then looked at the known physiology of the corresponding occipital-parietal-frontal cortex reach system (Fig. 2B) for inspiration to design coding schemes for the input and output layers of a feed-forward neural network (Fig. 2C). Finally, we used our black box model as a teacher to train the network to perform the 3D transformations for reach, much as the real system would learn through trial and error with sensory feedback. We compared the input and output properties of individual units within and between processing layers (Fig. 2D), using simulations of the major electrophysiological techniques (visual RF mapping, motor tuning, and microstimulation).

The overall purpose of this investigation was to 1) develop a theoretical network model for the feed-forward network properties that give rise to accurate visually guided 3D reach, 2) demonstrate through simulations how different “experimental techniques” can reveal different computational properties within this network, and 3) incorporate these findings, in light of previous models, into a single consistent theoretical framework. We show how our network performed the full reference frame transformation in a gradual manner through both serial transformations across successive hidden layers and through parallel distributed transformations across individual units within these layers. Gain fields are the necessary vehicle

for weighting the contributions of these units. We show that the neural populations, and even individual units, in these layers show different reference frames when tested using different techniques (Fig. 2D). Moreover, based on comparisons with the experimental data, we propose that this framework applies equally well to the physiology of the real system.

Materials and Methods

The visuomotor transformation process associated with visually guided reaching can be divided into 3 separate stages: 1) from the binocular 2D retinal images, the brain must construct and maintain an internal egocentric representation of the 3D location of the desired reach object and the initial hand position (Cohen and Andersen 2002; Merriam and Colby 2005; Tsutsui et al. 2005; Burgess 2006; Rushworth and Taylor 2006); 2) these egocentric, gaze-centered representations of the hand and target position then have to be transformed into a shoulder-centered movement plan for the hand (Burnod et al. 1999; Snyder 2000; Battaglia-Mayer et al. 2003; Crawford et al. 2004), and 3) the desired motor plan must be converted into dynamic muscle activation patterns that control the actual reaching movement (Kalaska et al. 1997; Baraduc et al. 2001; Todorov and Jordan 2002; Scott 2003). Here, we focus on the second step in this visuomotor conversion: how the brain performs the reference frame transformation from the egocentric, gaze-centered representations of hand and target position to the shoulder-centered reach movement plan. Because the motor command of the arm has to be specified with respect to its insertion point at the shoulder (Soechting et al. 1991) and only visual information (not proprioception) about the hand position was used, we modeled the visuomotor transformation between a gaze-centered and shoulder-centered motor plan and did not include the 3D geometry of the arm. The 3D arm geometry seems to be predominantly used in the third step to specify muscle activations from a desired movement plan (Kalaska et al. 1997; Kakei et al. 2001, 2003; Scott et al. 2001; Scott 2003).

Neural Network Model Architecture

We used a physiologically inspired, fully connected 4-layer feed-forward neural network to model the brain’s complete 3D visuomotor transformation for the planning of open-loop reach movements (Blohm and Crawford 2007). Figure 3 shows a schematic of the network

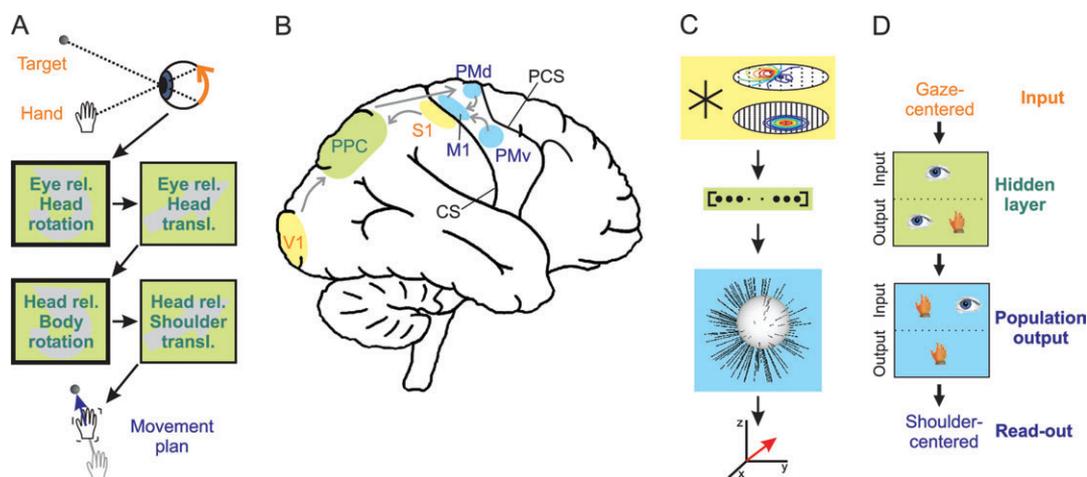


Figure 2. Different levels of understanding visuomotor transformations in the brain. (A) This geometric model illustrates the mathematical sequential transformations involved in converting a gaze-centered desired movement vector into a shoulder-centered motor plan. This transformation is composed by translations and rotations (Blohm and Crawford 2007). (B) Brain structures known to be part of the visuomotor transformation pathway in the brain. V1, visual cortex (gaze-centered encoding of viewed hand and target positions); PPC, presumably the hidden layer of our neural network model; S1, somatosensory cortex (a potential source of the extraretinal eye and head position signals); PMd/v, dorsal/ventral PM cortex (the hypothetical population output of our neural network model); M1, primary motor cortex; PCS, precentral sulcus; and CS, central sulcus. Panels (A, B) are adapted from Blohm and Crawford (2007). (C) Neural network implementation of the different brain structures. This is a cartoon of the neural network used in this study (see Fig. 3). (D) Our interpretation of how reference frame transformations might be performed in distributed computing. The eye icon stands for gaze-centered coordinates and the hand icon represents shoulder-centered coordinates. The presence of both icons depicts a spread of reference frames between and beyond gaze- and shoulder-centered coordinates. Same colors in panels (A-D) refer to corresponding levels of processing.

architecture. The first neural layer consisted of 7 distinct inputs, comprising retinal target and hand positions, the retinal disparity associated with these hand and target positions, 3D eye and head orientation inputs, and an ocular vergence input. We chose as a simplification to present initial hand position in visual coordinates and not to include any explicit proprioceptive signals because it has been shown that in the absence of vision, posterior parietal cortex (PPC) encodes hand position in visual coordinates (Buneo et al. 2002). This simplification is supported by the finding that the brain preferably uses visual input over proprioceptive information about hand position (Sober and Sabes 2003, 2005; Ren et al. 2006). As we will show, our findings concerning the hidden layer unit (HLU) RF properties are fully compatible with electrophysiological results (Buneo et al. 2002), which validates our approach.

All these inputs are necessary to fully describe the body geometry and to specify the 3D positions of hand and target in cyclopean eye-centered coordinates. The second (hidden) layer of our network was composed of a number of units that could vary between 9 and 100 units. The third (population output) layer contained a population of units that coded 3D movement plans in shoulder-centered coordinates. The activity of this layer was read out by the fourth (readout) layer, which coded the 3 components of the shoulder-centered movement plan in 3D euclidean space. All components of the network are explained in detail below.

The input-output relationship of all units in the second and third layers was modeled by a sigmoid function designed to mimic the nonlinear transfer function of real neurons (Naka and Rushton 1966a, 1966b, 1966c), that is,

$$a(x) = \frac{1}{1 + e^{-x}}. \quad (1)$$

The input layer activations were not put through this sigmoid function. The readout of the population coding in the output layer was purely linear (see below). Note that we did not use "basis function networks," as this has been done in previous studies (e.g., Pouget and Sejnowski 1997).

Inputs

Retinal Position: Topographical Hand and Target Maps

The sets of horizontal and vertical cyclopean (Ono and Barbeito 1982; Ono et al. 2002; Khokhotva et al. 2005) retinal positions (p_x, p_z) of hand and target were encoded in 2 separate retinotopic topographical maps of units, which specified hand and target direction relative to the fovea. These units had Gaussian RFs (width $\sigma = 20^\circ$), and their activations were computed as follows:

$$a_i = \exp\left(-\frac{(p_x - x_i)^2 + (p_z - z_i)^2}{2 \cdot \sigma^2}\right). \quad (2)$$

In analogy to the organization of the striate cortex, these neurons were uniformly distributed in a topographical map with a maximum circular eccentricity of 90° . Although we limited our visual inputs to 70° , we used 90° maximum eccentricity instead of the 70° visual field to avoid edge effects for the encoding of eccentric targets. The horizontal and vertical spacing of the units was 10° , which led to a total of 253 units. Similar topographical maps have been used to encode retinal target position in previous neural network studies (Zipser and Andersen 1988; Xing and Andersen 2000; Smith and Crawford 2005).

Retinal Disparity: Topographical Maps for Hand and Target

To specify hand and target distance, we encoded horizontal and vertical retinal disparities (d_h, d_v) of hand and target in 2 separate topographical maps of units. These units were given disparity tuning curves with profiles similar to those found in monkey neurons (Poggio and Fischer 1977; Poggio 1995) and cats (Nikara et al. 1968; Pettigrew et al. 1968; Ohzawa et al. 1997). The idealized disparity sensitivity functions we used here are 2D extensions of previously used ones (Lehky and Sejnowski 1990; Pouget and Sejnowski 1994). The activation of the topographical disparity neurons was computed as follows:

$$a_i = a_{X,i} \cdot a_{Z,i}, \quad (3)$$

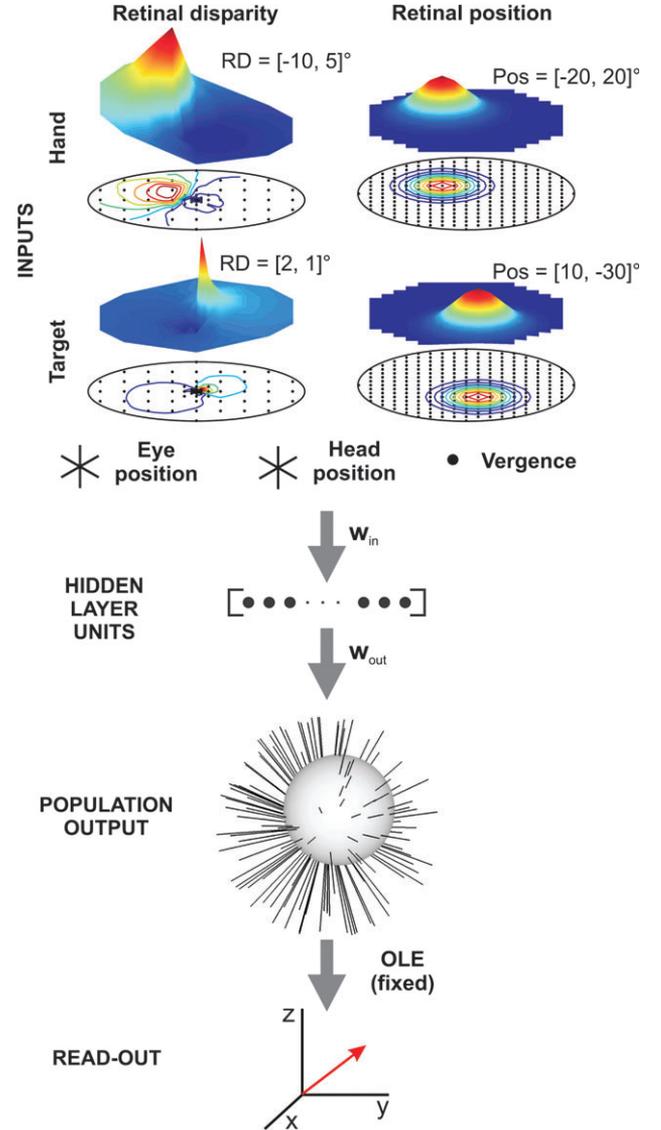


Figure 3. Neural network model. The 4 sections from top to bottom (separated by the downward gray arrows) show the 4 layers of our feed-forward artificial neural network. The input layer (first section labeled "inputs") is composed of 3D hand and target positions in gaze-centered coordinates, each of which is divided into 2 two-dimensional maps, that is, cyclopean retinal position and retinal disparity. In addition to these visual inputs, extraretinal 3D eye and head positions as well as an ocular vergence input were also provided. All signals were fed forward into a hidden layer that consisted of—depending on the training session—a different number of HLUs (9–100). The HLU signals were fed into a population output layer with 125 units with theoretical preferred movement directions that were randomly distributed in 3D shoulder-centered space. The spherical activation pattern superimposed on the preferred direction vectors designates the idealized cosine tuning of the population. The network was trained on the fourth layer, that is, the readout of the population code $\{x, \text{horizontal}; y, \text{depth}; z, \text{vertical}\}$. The weights of this readout were fixed prior to training and calculated using an OLE based on the population code's preferred directions and the assumption of cosine tuning. Weights w_{in} and w_{out} were the only weights that were adapted during training of the network. Refer to Materials and Methods for more details.

$$\text{where } a_{k,i} = \begin{cases} a_0 \cdot \exp\left(-\frac{(d_k - k_i)^2}{\sigma_i^2}\right) - a_1 \cdot \exp\left(-\frac{(d_k - k_i - \sigma_i^2)^2}{\sigma_i^2}\right), & \text{for } k_i \leq -1, \\ a_0 \cdot \exp\left(-\frac{(d_k - k_i)^2}{\sigma_i^2}\right) - a_1 \cdot \exp\left(-\frac{(d_k - k_i + \sigma_i^2)^2}{\sigma_i^2}\right), & \text{for } k_i \geq 1, \\ a_0 \cdot \exp\left(-\frac{(d_k - k_i)^2}{\sigma_i^2}\right), & \text{for } -1 < k_i < 1, \end{cases} \quad (4)$$

where k stands for X or Y . The variance was $\sigma_i^2 = (X_i^2 + Z_i^2)/2$ with a minimum of 10 minarc and we used constants $a_0 = 1$ and $a_1 = 0.35$

(Pouget and Sejnowski 1994). To maintain an approximately constant overlap of the disparity tuning curves, we used preferred disparities for our units (the disparity at which a unit showed maximum activation) that were nonuniformly spaced on an elliptic map. The spacing of these units was 10° horizontal and vertical for all preferred disparities with the exception of the central region (around zero). The spacing here was 1° for preferred disparities up to 2° horizontal and vertical and 5° spacing for units with preferred disparities up to 10°. The different combinations (X_i, Z_i) of the units' preferred disparity were arranged within an ellipse of (45°, 30°) and included a total of 67 units for each overall retinal disparity input signal. Again, the range of this ellipse was chosen slightly larger than the range of disparities used to avoid edge effects for the encoding of large retinal disparities.

Eye-in-Head, Head-on-Body, and Vergence Inputs

The 3D reference frame transformation depends critically on eye-in-head and head-on-shoulder positions (Blohm and Crawford 2007). Therefore, we need extraretinal signals that describe eye and head position. In addition, retinal disparity only provides distance relative to the fixation distance. To perform an accurate reach, we therefore need ocular vergence in order to obtain absolute distance.

For both eye-in-head and head-on-body orientations, we used a 3D angle vector representation (r_x, r_y, r_z), equal to the unit rotation vector multiplied by the rotation angle in degrees. We used an encoding scheme inspired by motor neuron activity. To encode positive and negative rotations (e.g., clockwise and counterclockwise), we transformed the 3D angle vector into a 6D array of input unit activities (we thus had 6 inputs for eye and 6 inputs for head position) arranged in push-pull antagonistic activations (King et al. 1981; Fukushima et al. 1990; 1992; Xing and Andersen 2000). Each pair of activations was computed as follows (Smith and Crawford 2005; Keith et al. 2007):

$$a_{\pm, i} = 0.5 \pm \frac{r_i}{2 \times r_0}, \quad (5)$$

where the maximum angle of orientation, r_0 , was equal to 50° for eye orientations and 70° for head orientations. The coordinate system for the encoding of the eye orientation angular vectors was rotated 45° around the z axis to reproduce the mixed vertical-torsional coding of eye orientations observed in the brainstem neural integrator (Crawford et al. 1991; Crawford and Vilis 1992; Crawford 1994; Suzuki et al. 1995).

We used a 1D (positive) input to code the ocular vergence angle of the eyes. Ocular vergence was defined as the absolute angle φ_V (in degrees) between the right eye and left eye gaze directions. Small angles correspond to far fixation positions, larger angles represent near fixation points. The activation of the input unit coding the vergence state of the eyes was computed as (Pouget and Sejnowski 1994):

$$a = \frac{\varphi_V}{45^\circ}. \quad (6)$$

Population Coding and Decoding of the Output

The output layer (fourth layer) of the neural network consisted of 3 units that coded movement in space. Each unit encoded a single spatial direction, that is, X (horizontal), Y (posterior-anterior), and Z (vertical) that corresponded to the movement distance of the hand along the 3 cardinal axes. These output units read out the distributed representation of the movement vector from the previous layer (third layer) of the neural network. This "behavioral" readout was chosen in a very specific manner that reflected the implicit assumption of cosine-tuned units in the population output layer (third layer) of our network. Note that the weights between layers 3 and 4 were calculated prior to the training and kept constant during the adaptation process of the neural network. We did not train the readout weights because this behavioral readout method was only used to quantify the movement vector encoded by the population output layer. As previously noted, decoding distributed representations is crucial because it allows an unambiguous quantitative interpretation of single-unit activity (Salinas and Abbott 1995).

The third layer of our neural network consisted of 125 cosine-tuned units with preferred directions (PD_i) randomly, uniformly distributed on a unit sphere (Fig. 3). Cosine-tuned neurons that encode movement

direction in extrinsic (likely shoulder centered) coordinates have been observed in the premotor (PM) cortex of the monkey (Kalaska et al. 1997; Kakei et al., 2001, 2003; Scott 2001). It has also been shown theoretically that cosine tuning was optimal for motor control in 3D (Flash and Sejnowski 2001; Todorov 2002). To obtain such a spherically uniform random distribution of preferred directions, we generated 3 random Gaussian variables (x_i, y_i, z_i) with a mean of zero and a standard deviation of one. Next, the distribution of the preferred direction vectors was computed as:

$$\vec{PD}_i = \frac{1}{\sqrt{x_i^2 + y_i^2 + z_i^2}} \begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix}. \quad (7)$$

\vec{PD}_i is statistically uniform over the spherical surface (Muller, 1959; Marsaglia 1972). We used a statistically uniform distribution of PD_i to match the above-cited electrophysiological findings. In order to calculate the behavioral readout weights, we assumed cosine tuning in the third (population output) layer units. The hypothesized cosine-tuning behavior of each third layer unit i can then be represented by the units' theoretical activation as:

$$a_i = a_0 + a_1 \cdot \cos \theta_i, \quad (8)$$

where $a_0 = 0.5$ is the baseline firing rate and $a_1 = \frac{0.5 \cdot \|\vec{M}\|}{A_{\max}}$ is the cosine scaling parameter that scaled the unit activity to the size of the required movement \vec{M} , whereas the angle $\theta_i = \cos^{-1} \left(\frac{PD_i \cdot \vec{M}}{\|\vec{M}\|} \right)$ coded the direction of the movement. The combined coding of movement direction and amplitude within the same units is consistent with recent neurophysiological findings (Fu et al. 1993; Messier and Kalaska, 2000). The maximum amplitude of movement was $A_{\max} = 2m$, and the activation of the output units was scaled so that $a_i \in [0; 1]$. The maximum movement amplitude A_{\max} was larger than the maximal possible movement (=175 cm). This was done to avoid the case where the desired output of the units was close to the working limit of the sigmoid transfer functions. Note that we did not train the network to reproduce the theoretical activations specified in Equation (8).

Importantly, the implicit assumption of a cosine-tuning behavior of the third layer units allowed us to explicitly compute the readout weights from these third layer units i into the units j in the output (fourth) layer. To do so, we used an optimal linear estimator (OLE) method (Salinas and Abbott 1994). Using this method, we calculated the weight matrix w_{ij} between layer 3 and layer 4 (which is also called the "OLE") as:

$$w_{ij} = \sum_k Q_{ik}^{-1} L_{kj}. \quad (9)$$

For a full cosine-tuning function of the third layer units as described in Equation (8), the center of mass matrix L_{kj} (the index j stands for the vector component, i.e., X, Y , or Z) and the cross-correlation matrix Q_{ik} were calculated as follows:

$$L_{kj} = \int dM_j \cdot M_j \cdot a_k(M_j) = \frac{4\pi \cdot a_1}{15} PD_{kj}, \quad (10)$$

$$Q_{ik} = \sigma_k^2 \delta_{ik} + \int d\vec{M} \cdot a_i(\vec{M}) \cdot a_k(\vec{M}) \\ = \sigma_k^2 \delta_{ik} + \frac{4\pi \cdot a_0^2}{3} - \frac{4\pi \cdot a_1^2}{15} \cdot \left(\vec{PD}_i \cdot \vec{PD}_k \right). \quad (11)$$

The cross-correlation matrix Q_{ik} includes an estimate of the expected neural noise σ_k and a dot product that specifies the interaction between 2 tuning curves. We set the noise parameter at an arbitrary value of $\sigma_k = 0.01$, which was constant across all third layer units k . See Supplementary Methods for a description of the theoretical readout accuracy of the movement vector for different noise levels. We chose the ideal number of third layer units based on the observation that improvement in accuracy was small when the number of units increased past 125. Again, the readout weights between the third layer, and the output layer were assigned prior to the network training and were not modified during the training process. We trained our neural network on the output layer (the 3D movement vector) only and did not constrain the activations of the units in the third (cosine tuned) population output layer. It is also important to note that the choice of a uniform distribution of the third layer PD_i did not affect or constrain

the readout process in any way because the OLE does not require any particular distribution of the preferred directions.

Training Method and Training Set

We generated a training set that accounted for the complete 3D geometry of the eye-head-shoulder linkage (Blohm and Crawford 2007). Within this training set, the 3D binocular eye positions complied with the binocular extension of Listing's law (Van Rijn and Van den Berg 1993; Hepp 1995; Tweed 1997; Somani et al. 1998), which constrains the 3 degrees of freedom (df) for the eye rotation behaviorally to 2 effective df. This places the eye rotation vectors into a plane known as Listing's plane. The binocular version of Listing's law is modulated by the static vestibuloocular reflex (VOR) by counter-rolling the eyes when the head is tilted toward the shoulder (ocular counter-roll) and by modifying the primary position of Listing's plane with head up and down (gravity pitch of Listing's law) movements (Haslwanter et al. 1992; Bockisch and Haslwanter 2001).

Eye and head orientations were randomly chosen and were approximately uniformly distributed around straight-ahead position. Fixation distance varied between 25 cm and 5 m so that vergence was approximately uniformly distributed. We then randomly chose a combination of hand and target positions within the visual field, that is, at a maximum of 70° visual eccentricity. The range of both hand and target positions was set within reach space, that is, not more than 85 cm distant from the right shoulder (here, we arbitrarily chose to simulate right-hand motor planning).

From this visuomotor arrangement, we computed the projections of hand and target onto retinal coordinates of a hypothetical cyclopean eye. We also calculated retinal hand and target disparity, eye position, head position, ocular vergence, and the resulting motor plan in shoulder-centered coordinates. We randomly generated a total of 500 000 training points, where each training point corresponded to one set of input and output activations computed for one particular eye-head-hand-target configuration. A random subset of this training set was used to train our networks (Table 1 shows the size of the training set for different network sizes). See Results for more details.

We used a resilient back-propagation (RPROP) technique to adjust the weights of the neural network during training (Riedmiller and Braun 1993). As a modification of the pure gradient descent algorithm, RPROP dynamically modifies the learning rate as a function of the gradient's sign but independently of the size of the derivative. This results in an efficient adaptation process with increased convergence and in a stable learning behavior. Note again that only the interlayer weights of layers 1-2 and 2-3 were adapted. The weights between layers 3 and 4 for the readout of the cosine-tuned activity were held constant.

The neural network was implemented in Matlab 7 (R14) (Mathworks Inc., Natick, MA) using the Neural Networks Toolbox and customized functions. We used a 64-bit dual Intel Xeon Irwindale (3.0 GHz, 800 MHz system bus, 2 MB integrated L2 cache) computer with 8 GB RAM (400 MHz DDR2) and running a RedHat Linux Enterprise 4 operating system. Training durations varied from a few hours (9-HLU network) to approximately 4 weeks (100-HLU network) and depended on the criterion of convergence as well as the size of the training set (Table 1). We stopped network training when the evolution of the root-

mean-squared error (RMSE) was no longer perceptible on a log-log scale, that is, the gradient became $<10^{-6}$.

Neural Network Analysis

To analyze the network, we used methods similar to those employed earlier in oculomotor models (Smith and Crawford 2005; Keith et al. 2007). We quantified the overall network performance by computing the 3D compensation index (Blohm and Crawford 2007). Briefly, the 3D compensation index is a metric measurement that assesses the amount by which the network adjusted the gaze-centered movement vector to produce the shoulder-centered motor command.

We also computed eye and head position sensitivity vectors (Keith et al. 2007). These are 3D vectors that describe how the activity of a certain HLU or unit of the population code (third layer) is modulated by a change in 3D eye or head position. For example, a purely horizontal eye position sensitivity vector would indicate that only horizontal eye position changes modulate the unit's activity, but the activity remains constant across vertical or torsional eye movements. The sensitivity vectors are defined by the weights connecting the eye or head position input to the unit considered.

We computed motor fields in order to assess a unit's contribution to the motor output. To analyze the motor fields, we fitted the unit activity for all movements executed with a certain eye position to the following generalized cosine tuning function.

$$a_i = b_i + c_i \cdot \cos\theta_i. \quad (12)$$

The index i stands for an individual movement. The parameters were calculated using the following equations.

$$\begin{aligned} b_i &= b_0 + b_1 \cdot \|\vec{M}_i\|, \\ c_i &= c_0 + c_1 \cdot \|\vec{M}_i\|, \\ \theta_i &= \cos^{-1} \left(\frac{PD \cdot \vec{M}_i}{\|\vec{M}_i\|} \right). \end{aligned} \quad (13)$$

We used a nonlinear least-squares fitting algorithm (Gauss-Newton search) to evaluate the free parameters b_0 , b_1 , c_0 , c_1 , and PD for each eye position. Next, we computed the rotational gains to evaluate the change of the preferred direction PD with eye position, that is, how eye position changes the direction of movement for which we get maximum activity for a given unit. To do so, we calculated the angles between the preferred direction for nonzero eye positions (i.e., eye positions that are not straight ahead) and the preferred direction for straight-ahead fixation. To obtain the rotational gains, we then performed a linear regression of those angles with the amplitude of eye position. We used rotational gains as one way of quantifying the motor reference frame of each unit. Next, we computed the gains related to the change in motor field amplitude (and not direction) with eye position. We computed the unit's activation at the preferred direction for each eye position using the identified generalized tuning parameters in Equations (12) and (13). We then performed a regression analysis of the unit's preferred activation as a function of eye position for different eye positions, which resulted in the motor field amplitude change gain value. (Note: we multiplied this gain by the eye position range, i.e., 90°, to render the result dimensionless.)

Finally, we computed response field gradients, which provide an indicator of which variables modulate the unit's response most strongly. To do this, we varied target position, fixation position, and initial hand position separately in 5° steps from -45° to 45° horizontally. All 3 positions were in a frontoparallel 50-cm distant (from the eyes) tangential plane as this was the case for electrophysiological experiments (Buneo et al. 2002; Pesaran et al. 2006). We calculated the HLU (second) and population output (third) layer unit activity for each combination of eye, hand, and target position (e.g., Fig. 11A,B). We then computed the gradients of unit activity across all positions and for all 3 pairs of possible combinations of eye, hand, and target positions. For example, the gradients in Figure 11A are generally directed downward, that is, along the greatest rate of change. Doing this for all 3 combinations of eye, hand, and target position resulted in 3 gradient fields that could be represented as local rates of change in unit activity for each eye/hand/target position (we calculated the gradient at each pixel of Fig. 11A,B). In order to extract a single index for the encoding

Table 1

Global network performance evaluated on $N = 10\,000$ test points

# HLU	3D compensation		Network error (cm)		RMSE	# Training points
	R^2	Slope	Mean	Standard deviation		
9	0.878	0.873	14.20	8.43	0.0417	15 000
16	0.936	0.934	10.06	6.58	0.0303	25 000
25	0.960	0.953	7.86	5.36	0.0249	45 000
36	0.972	0.963	6.44	4.75	0.0206	60 000
49	0.976	0.971	5.96	4.25	0.0192	80 000
64	0.982	0.975	5.23	4.05	0.0173	110 000
81	0.984	0.980	4.81	3.60	0.0162	140 000
100	0.992	0.989	3.25	2.73	0.0114	170 000

scheme of each unit, we multiplied the direction of each gradient by a factor 2 and then averaged across all gradient vector directions within the gradient field. The direction of the resultant vector was then used as an index indicating which encoding scheme was used, that is, either encoding individual variables or encoding combinations of variables, such as eye + hand or eye - hand (see also Buneo et al. 2002).

Results

Network Performance

Before analyzing the “neural mechanisms” within an artificial network, it was first necessary to confirm that the network has learned the relevant aspects of the task. The details of the 3D visuomotor transformation of reaching are geometrically complex and highly nonlinear, so we summarized the overall performance of the network across all geometrical configurations in the following ways. We will consider an example network with 36 HLUs. Figure 4A shows a histogram of absolute reach errors produced by the network for 10 000 arbitrary eye and head positions. As can be observed, the large majority of absolute reach errors were smaller than 10 cm (mean = 6.4 cm), which is similar to human behavior (e.g., Blohm and Crawford 2007). This confirms the good average performance of the network.

We characterized the overall performance of all networks in Table 1, as a function of the number of units in the second (hidden) layer. As expected, the more second layer units there were, the better the performance. This can be seen in the RMSE value, which compares the desired to the observed activations of the output. A more intuitive indicator is the network error, where we indicated the mean reaching error in centimeters produced by the neural network for a random subset of 10 000 test points that were generated in the same manner as the training points (see Materials and Methods).

A quantitative analysis was performed using the 3D compensation index, which assesses how well extraretinal signals were taken into account in the visuomotor transformation (see Materials and Methods). For example, Figure 4B shows the 3D compensation produced by a typical 36-HLU network, computed as a function of the predicted (optimal) 3D compensation. The network produced observed 3D compensation values closely matching the 3D compensation predicted by our analytical model (Blohm and Crawford 2007). This can be observed when considering the value of the slope between observed and predicted 3D compensation values, which in the case of a 36-HLU network (Fig. 4B) was 0.963. As can be seen for all trained networks in Table 1, this slope provides the mean percentage of compensation of eye/head orientations, and the R^2 value gives an indication of the linear goodness of fit of the data, that is, the fraction of variance accounted for. If the slope was zero, the extraretinal signals would not have been taken into account and the network would perform reaching movements as though the eyes and head were straight ahead, which would then produce large errors (Blohm and Crawford 2007). On the other hand, a slope of 1 would indicate that the network on average fully accounted for the linkage geometry of the body and performed accurately, within a precision expressed by the R^2 value. Overall, Table 1 shows that all networks performed reasonably well, with performance improving along with the number of HLUs. In particular, the performance of the 36-HLU networks was quantitatively similar to that observed in real human subjects (Blohm and Crawford

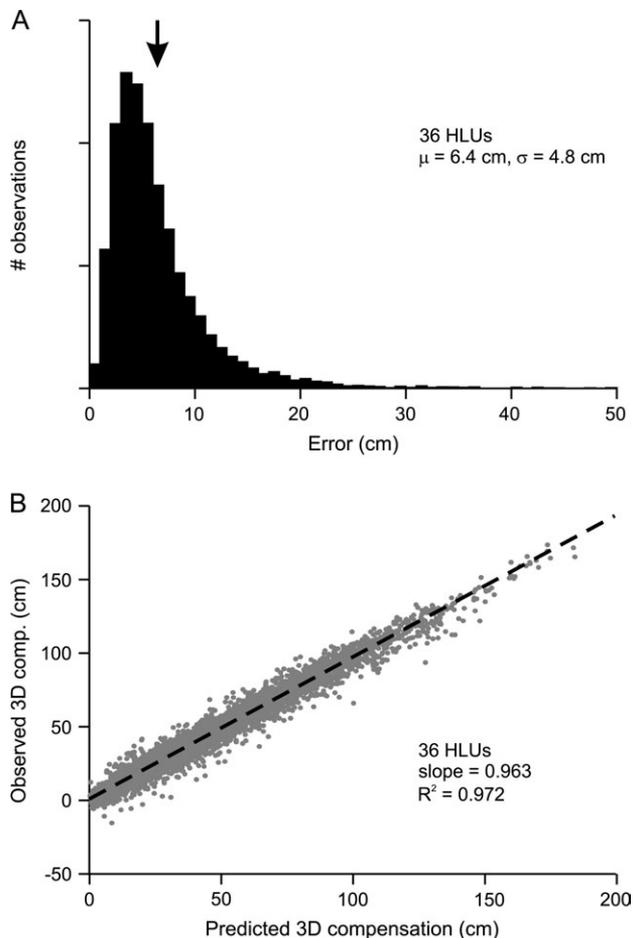


Figure 4. Performance of a typical 36-HLU (second) network. (A) Histogram of reaching errors produced by the 36-HLU network for a random test set of arbitrary eye-head, hand, and target positions. The mean absolute reaching error was 6.4 ± 4.7 cm (mean reach magnitude = 53.8 cm). (B) Observed 3D compensation index as a function of the ideal, predicted 3D compensation index. The 36-HLU network accounted on average for 96% (see slope) of the extraretinal eye and head position signals to adjust the gaze-centered motor plan to produce a correct shoulder-centered movement.

2007). In the following sections, we will describe the network behavior showing typical examples of the 36-HLU network (because it performed similarly to human subjects) and we will provide population results across all networks.

Network Analysis: General Considerations

The goal of this paper was to investigate the mechanism by which reference frame transformations could be achieved with distributed processing and to make predictions about expected neural properties in brain areas involved in this process. To answer the first part of this question in a way that is directly relevant for neurophysiological studies, we assessed how individual units in specific layers of the neural network transform information through their input-output relationships. To identify reference frames, we analyzed whether a unit's activity was modulated with eye or head position. For example, a unit's preferred direction for encoding the visual input in gaze-centered coordinates would be independent of eye and head position, whereas the preferred direction would shift if the units used some other reference frame.

To investigate the individual units' input-output relationships, we chose to perform an analysis that was inspired by

neurophysiological techniques. To obtain the input reference frame of a unit, we aligned the unit's activity with the visual input (visual RF) and investigated how this visual RF was modulated with eye and head position. To obtain the output reference frame, 2 neurophysiological techniques were used: 1) alignment of the unit's activity with the motor output (motor field) and 2) using simulated microstimulation, which sets the value of an individual unit's activity artificially to the maximum and looks at the effect this has on the motor output. Using these 2 methods, we investigated whether the motor fields or simulated microstimulation results were modulated by different eye/head positions. We will provide more details about each individual technique when describing them hereafter.

Input Properties: Visual RFs

To begin, we investigated the input properties of the individual units of the second (hidden) layer. To do this, we computed the visual RF for each unit in this layer, which we did by holding all hand-related inputs, as well as target retinal disparity, vergence, eye, and head positions, constant. (Note: because the encoding of initial hand and target positions are strictly the same, all findings apply for both variables and we only show the results of changing target position.) We then presented targets at all possible horizontal and vertical visual locations and computed the resulting activations of the HLUs. Figure 5A–D shows examples of 4 typical second (hidden) layer units' RFs. We represented each unit's activity by means of a color code for each location within the 90° visual field. In Figure 5A, for example, a target presented in the lower visual field would activate this particular unit, whereas a target presented in the upper visual field would result in very little activation. Therefore, this particular unit's visual RF is in the lower visual field. This is indicated by the pink, black-bordered square that shows the location of the center of mass of the RF.

Up to this point, one cannot make any conclusions about the reference frame in which these units encode incoming visual information. To do so, one has to change eye and/or head position and investigate whether the RFs change their preferred location—that is, does the center of mass shift with eye/head position? (Note: here, we only illustrate changes in eye position. Head position is encoded in the same way as eye position and provided qualitatively similar results.) To examine the influence of eye position on the RF, we first plotted the sensitivity vector (horizontal and vertical components only) in the RF plots as black bars (Fig. 5A–E, see Materials and Methods). The sensitivity vectors represent the direction in which the eyes have to move in order to maximally modulate the unit's activity, and its size is proportional to the strength of this modulation. Therefore, they indicate in which directions we have to move the eyes in order to analyze the effect of eye position on a second (hidden) layer unit's activity.

We can make 2 predictions of what we might expect to find: 1) if the RF encodes visual information in gaze-centered coordinates, that is, it is only important where targets are relative to the line of sight but not where they are in space, then the center of mass should be independent of eye position or 2) if the visual RF encodes targets in shoulder-centered coordinates, then the center of mass should shift in the direction opposite of the eye orientation in order to maintain a spatially stable code in body-centered coordinates.

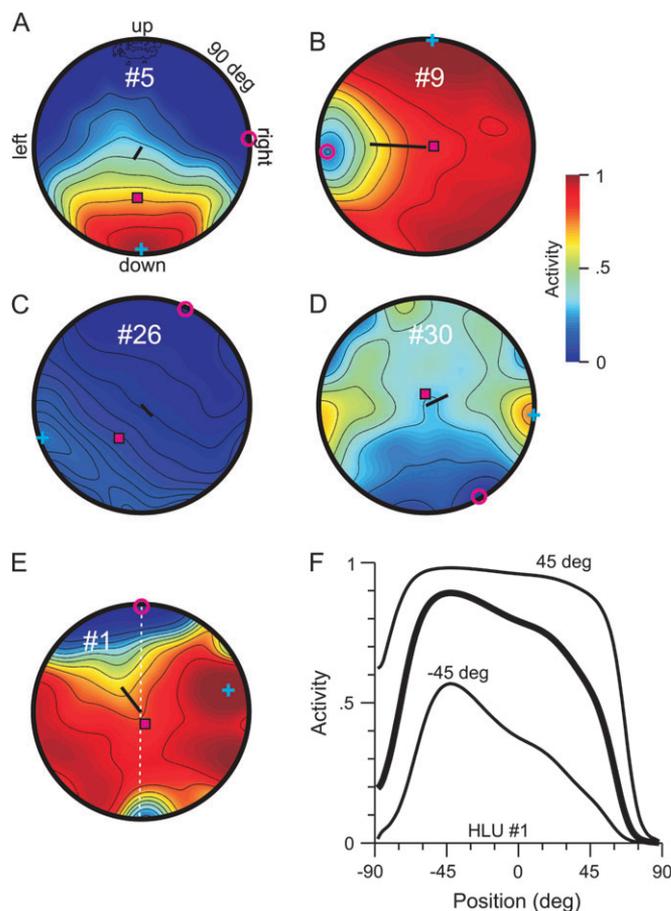


Figure 5. Visual RFs in the hidden layer. Panels (A–E) show typical RFs from 5 HLUs from a 36-HLU network. The activity (color coded) of each unit is plotted as a function of visual target position ($\pm 90^\circ$ horizontal and vertical range), whereas all other inputs were kept constant, that is, central hand position, straight-ahead eye, and head position and constant vergence. Landmarks such as the location of the maximum (cyan cross), minimum (magenta circle), and the center of mass (black and magenta square) of the unit's activity are indicated. We also plotted the horizontal-vertical eye position sensitivity vector (black bar), which indicates that the direction the eyes would need to move to maximally influence the unit's activation. The relative size of these vectors indicates the relative amplitude of change with eye position. Thin black lines represent isoactivity lines. (F) Activity of HLU #1 (panel E) along a cut through the center of the RF and the position of the minimum (dotted white line in panel E). Different lines indicate activity for different vertical eye positions between -45° (down), 0° , and 45° (up). The bold line represents the 0° eye position.

We examined the influence of eye position on a typical second (hidden) layer unit's activity and plotted the activity of the example shown in Figure 5E for different eye positions. For easier comparison, we do not show the complete RF but only a 2D slice through the minimum (indicated by the magenta circle in Fig. 5E) and the center of the visual field (dotted white line). The activity for straight-ahead eye position (as in Fig. 5E) corresponds to the bold line in Figure 5F and shows a hill-like pattern. As can be observed, changing eye position from -45° to 45° essentially gain modulates the unit's activity (the activity moves up and down) but does not much shift its location (left or right in position), similar to the so-called gain field mechanisms that have been observed in most parts of the cortex involved in visuomotor transformations (Andersen et al. 1985; Zipser and Andersen 1988; Salinas and Abbott 1995). This is the same for all second (hidden) layer units (shown later in Fig. 7). Note that the changes in the shape of the RF across

different eye orientations in Figure 7E were due to the saturation of the sigmoid transfer function. Because there is almost no shift in these unit's RF locations (i.e., the centre of mass did not change with eye position), we will interpret this as a gaze-centered encoding scheme.

We also examined the input reference frame in the third layer, that is, the population code of the desired movement vector, by again investigating the influence of eye position on the RFs of this layer. To do this, we considered how the RF varied with horizontal and vertical eye position in an example unit. This is shown in Figure 6A-I for a typical third layer unit, #18. Panel (E) shows the visual RF of this unit for a straight-ahead eye position. As can be seen in panel (F), if the eyes move 40° to the right, the RF shifts to the left. This can be seen by observing the change in the position of the centre of mass (pink, black-bordered square). Likewise, if the eyes rotate 40° to the left, the RF shifts to the right (panel D). Similar behavior can be observed for vertical (panels B and H) and consequently also for oblique eye positions (panels A, C, G, and I).

To obtain the entire representation of the RF shift for different horizontal and vertical eye positions, we changed eye position in a more systematic fashion, that is, in 5° steps independently for the horizontal and vertical directions. For every eye position, we computed the horizontal and vertical position of the RF, quantified by the center of mass position. We then plotted the relative positions of the center of mass positions in Figure 6J for the example unit #18 shown in panels (A-I). Each dot represents one center of mass position, and dots from adjacent eye positions are connected through the solid line. Thus, the intersection of the horizontal line corresponds to the straight-ahead eye position example of panel (E). The center of mass moves to the right (left) when the eyes move left (right) and up (down) when the eyes move down (up). Clearly, this unit seems to shift its RF toward maintaining a spatially stable representation of the visual object. Therefore, we conclude that this unit uses an input code that approaches shoulder-centered coordinates.

In the next step, we further quantified the RF shift to perform a more formal reference frame analysis. We performed a regression analysis on the center of mass shift (as shown in Fig. 6J) as a function of horizontal and vertical eye position for each unit in the neural network. This regression analysis provided a gain factor indicating the extent to which eye position modulated the position of the center of mass. If there is no center of mass shift, then the gain factor is zero, indicating gaze-centered coding. If the gain factor is -1, then the RF shifts in the opposite direction and by the same amount as the eye orientation and, thus, maintains a spatially stable representation, that is, codes positions in shoulder-centered coordinates.

The result of this analysis is shown in Figure 7, where we plotted the horizontal and vertical centre of mass gains for horizontal (Fig. 7A) and vertical (Fig. 7B) eye movements. Each dot represents the behavior of one unit of the third layer of our 36-HLU neural network, that is, the population output. For horizontal eye movements, we observed a large distribution of horizontal shift gains and a narrower distribution of vertical shift gains (see histograms on the axes). In contrast, vertical eye movements (Fig. 7B) resulted in a narrower horizontal gain distribution and a broader vertical center of mass shift gain distribution. This means that eye movements mainly result in

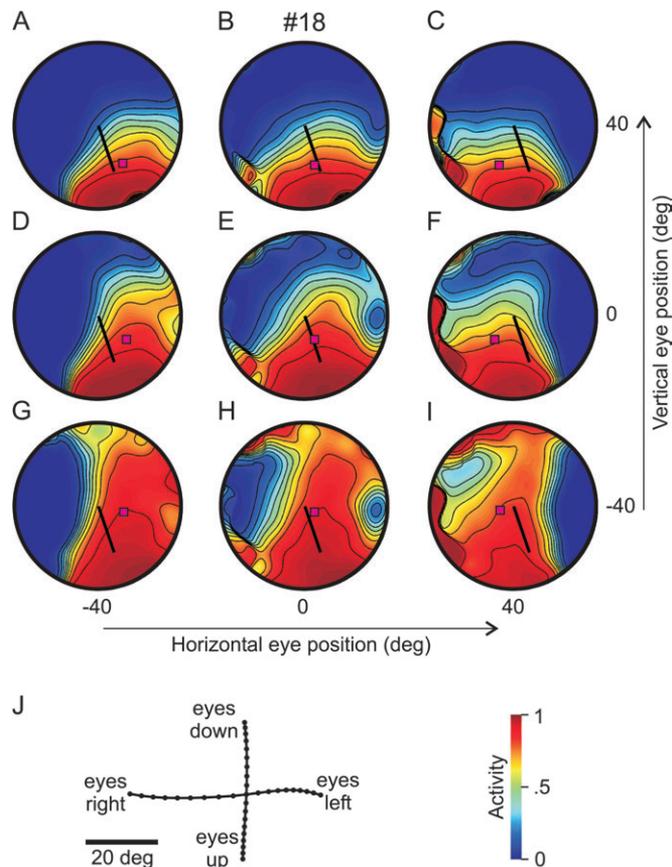


Figure 6. Visual RF modulation of population code unit #18 as a function of eye position. Panels (A-I) show RFs of the same unit plotted in the same manner as in Figure 5 but for different horizontal and vertical eye positions. Magenta and black squares, center of mass; black bar, motor sensitivity vector, that is, the movement direction toward which this unit contributes most (only horizontal and vertical component). Eye positions in degrees were (horizontal, vertical) (A) (-40, 40), (B) (0, 40), (C) (40, 40), (D) (-40, 0), (E) (0, 0), (F) (40, 0), (G) (-40, -40), (H) (0, -40), (I) (40, -40). (J) Representation of the horizontal and vertical shift of the center of mass as a function of horizontal and vertical eye positions in 5° steps (black dots). For this unit, for example, if the eyes moved rightward, the center of mass moved leftward (left is left and up is up).

shifts parallel to the eye movement but also modulate the visual RFs in the orthogonal direction to a smaller degree. For example, a horizontal eye movement can also evoke vertical RF shifts, although the RF shift is predominantly horizontal at the population level.

To quantify the overall range of shift gains, we used the horizontal gain related to the horizontal eye position change and plotted it as a function of the vertical shift gain evoked by the vertical eye position change (Fig. 7C). The gray box indicates the range of the observed gains across all third layer (population output) units, and the cross shows the mean value for this 36-HLU network. In Figure 7D, we show the range of gains represented in the same way as in panel (C) for all network sizes. We observed a broad distribution of gain values in all networks as shown by the histograms. A purely gaze-centered unit would have horizontal and vertical gain values of zero, whereas a shoulder-centered coordinate frame would result in gain values of -1. We interpret this large range of gain values as reflecting different units whose input sensitivity is not fixed with respect to one particular reference frame but rather

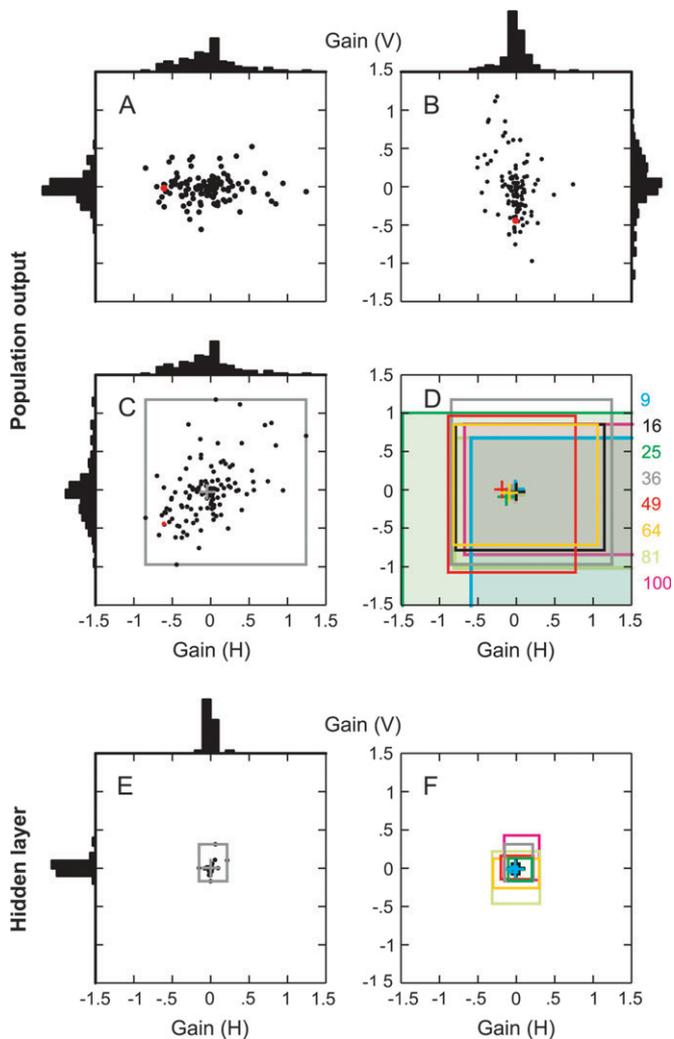


Figure 7. Reference frame analysis of the hidden layer and population code using RFs. (A) RF shift gains (horizontal and vertical) for horizontal eye movements. Each dot represents the gain of one unit in the population code of a 36-HLU network. The red dot depicts unit #18 analyzed in Figure 6. Zero gain values indicate gaze-centered coding; a horizontal gain value of -1 would classically be interpreted as shoulder-centered (or space fixed) coding. (B) Same representation for vertical eye movements. Here, a vertical gain value of -1 would be shoulder-centered coding. (C) Combined horizontal and vertical shift gain for the same population code. The vertical gain for vertical eye positions is plotted as a function of the horizontal gain for horizontal eye movements in order to capture the main properties of the RF shifts. The gray square indicates the range of obtained gain values for this 36-HLU network and the gray cross depicts the mean gain. (D) Summary of gain values through the population codes of all networks. The colored numbers shown to the right of the graph indicate the network sizes. Squares that go beyond the limits of the graph are semitransparent. In all network sizes, there is a wide variety of gain values present for the population code layer. (E–F) The same analysis for the HLUs for comparison. Panel (E) shows the details for the example 36-HLU network and panel (F) summarizes the behavior of all networks.

is weighted between gaze-centered and shoulder-centered coordinates. This has sometimes been called an “intermediate reference frame” (e.g., Buneo and Andersen 2006). These results can be compared with the results of the same analysis performed on the second (hidden) layer of the 36-HLU network (Fig. 7E) and for the different sizes of neural networks used in this study (Fig. 7F). This confirms the findings from Figure 5 showing only close to gaze-centered reference frames for the visual RFs.

Output Properties: Motor Fields

Up to this point in the analysis, we have analyzed the input reference frame for each unit in the hidden layer (second layer) and the population output layer (third layer). In general, one tends to assume that a properly tuned visuomotor network should contain units whose visual and motor tuning is aligned (so that vision results in corresponding movement), but in networks involved in coordinate frame transformations, there is good reason to believe that neural populations and even individual units should deviate from this scheme (Pellionisz and Llinas 1985; Pellionisz and Ramos 1993; Crawford and Guitton 1997; Pouget and Sejnowski 1997; Pouget and Snyder 2000; Smith and Crawford 2005). In particular, in 3D reference frame transformations, visual input in eye coordinates misaligns with the behavioral output in shoulder coordinates as a function of the orientation of the sensor relative to the shoulder (Klier and Crawford 1998; Crawford et al. 2000), and there should be an underlying neural mechanism to account for this. Because our network was trained to perform such a 3D transformation (Blohm and Crawford 2007), we hypothesized that the hidden layers of the network would show different input and output properties, even at the level of individual units. To test this, we investigated 2 output properties of the neural network units. We first considered motor fields. As opposed to visual RFs, where the activity of a unit is correlated with the visual input, in a motor field, the unit activity is instead correlated with the 3D movement direction in space.

Motor fields thus provide information about the motor output of a unit, that is, how a unit’s activity changes as a function of the movement produced. To compute motor fields, we have to produce movements covering all 3 dimensions of space and measure a unit’s activity related to those specific movements. To accomplish this, we will align the unit’s activity with the produced motor output (instead of aligning it with the visual input as for the visual RFs). If a unit preferentially participates in generating movements directed to a specific location in space, we expect a unit to display a preferred direction, meaning that it would discharge most when movements are oriented to that certain portion of 3D space.

Let us first consider a typical example motor field from HLU #17 of our 36 HLU network. Because 3D motor fields are difficult to represent graphically, we will show in Figure 8A–C, a 2D cut through the direction of maximal activity (measured for straight-ahead eye and head position) in order to demonstrate directional and amplitude tuning of motor fields with eye position. We will then proceed in the same manner as for the visual RF analysis and change eye/head position to see how the motor field changes. Figure 8 shows the motor field for 40° leftward eye position (panel A), straight-ahead eye position (panel B), and 40° rightward eye position (panel C). The motor field can be seen to change for different eye positions. Indeed, eye position affected both the preferred direction of the motor field (red lines), which rotated in the direction of the eye movement, and the amplitude of the motor field in a gain-like fashion, that is, the cosine tuning became smaller in amplitude when the eyes moved rightward. Because the preferred direction (in spatial coordinates) shifted with eye position, it means that the movement vector was approximately constant relative to gaze. Indeed, the preferred direction of the motor field rotated by 69.6° for a 80° total eye orientation change.

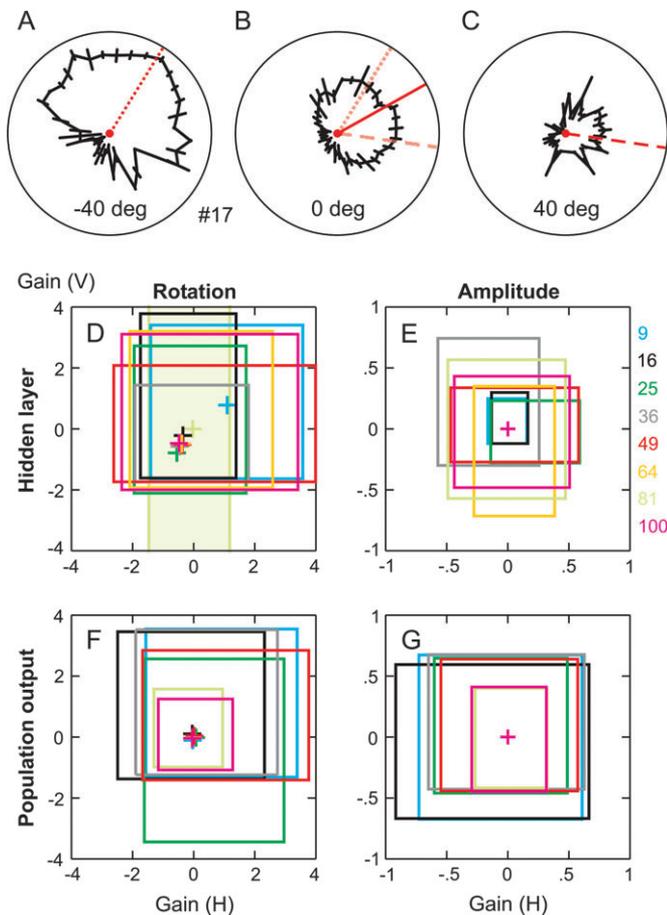


Figure 8. Motor field reference frame analysis for hidden layer and population code. (A–C) A 2D cut through a typical HLU motor field (unit #17) in the 36-HLU network (normal vector of plane: $[0.14; -0.29; 0.95]$). The unit’s activity is plotted in polar coordinates (maximum activity = 1) as a function of movement direction relative to the theoretical preferred direction of this unit (mean and standard deviation, see text for details). (A) Eyes at -40° (to the left). (B) Eyes at primary position (0°). (C) Eyes at 40° (to the right). The red lines indicate the reconstructed tuning direction from the cosine function fit. The solid red line corresponds to the tuning curve in panel (B). The dotted and dashed lines stand for the tuning curves in panels (A, C) and are superimposed on panel (B) for clarity. The different red lines show that this unit’s tuning direction and amplitude change with eye position. (D) Quantification of the rotational gain modulation of the tuning with horizontal and vertical eye position (same representation as Fig. 7D,F) for the hidden layer for all networks. (E) Amplitude gain modulation of the hidden layer for all networks with eye position. (F) Rotational gains for the population code layer across all networks. (G) Amplitude gain ranges for the population code layer for all networks.

Therefore, we interpret this unit’s motor field as displaying a reference frame close to gaze-centered coordinates. In addition, the change in amplitude gain values with different eye positions suggests eye/head position gain modulation.

We evaluated the influence of eye position on the motor field direction and amplitude for each unit in the hidden (second) and population output (third) layer (see Materials and Methods for details of our calculations). The procedure was similar to that used for the visual RFs. Figure 8D shows the ranges of the rotational gain for the hidden (second) layer of all different network sizes, presented in the same way as in Figure 7D,F. We observed a large range of rotational gain values that often extended to values larger than 1 (a gain value of 4 means that from 45° leftward to 45° rightward eye position the motor field

would have performed one complete revolution, i.e., a 90° eye position change would result in a 360° preferred direction change). Similarly, we also observed a large range of “amplitude” gain values, shown in Figure 8E. Here, a gain of 1 means that the motor field is modulated maximally across the complete range of possible values (between 0 and 1). In contrast, a gain of 0 means that eye position does not affect the amplitude of the motor field. The sign of the amplitude gain was chosen such that a positive gain meant an increase in amplitude for an upward or rightward eye orientation. Overall, the HLUs (second layer) show large amplitude gain modulation of the motor fields (Fig. 8E). In addition, the neural network uses a mixture of any imaginable reference frame in the individual HLUs (second layer) in terms of the unit’s contribution to the motor output, that is, gaze-centered, shoulder-centered, and reference frames intermediate between gaze and shoulder centered (Fig. 8D). It is difficult to interpret these large rotational gains in terms of reference frames. As a general observation, the networks used the complete parameter space to flexibly obtain the required behavior and combined these different representations at the population level in a purposeful fashion by making use of the amplitude gain modulations.

We observed broadly distributed behavior for the motor field reference frame analysis of the population output (third) layer units similar to what we have found in the hidden (second) layer. This is shown in Figure 8F for the rotational gains indicating the preferred direction change with different eye positions and is also present in Figure 8G for the amplitude gains describing the scaling of the motor fields with changes in eye position. Although both layers showed cosine-tuned properties, they did not use any particular identifiable reference frame. This is particularly surprising for the population output (third) layer that was (indirectly) designed to encode movement vectors in shoulder-centered coordinates (see Discussion). Therefore, in the classical view of motor fields addressing the output reference frame, one would have expected to find purely shoulder-centered coordinates.

Output Properties: Simulated Microstimulation

Another method to assess the output properties of individual units is to simulate microstimulation in the network. Microstimulation primarily modifies the neural network downstream of the locus of stimulation (Pare and Wurtz 1997; Tehovnik et al. 2003; Smith and Crawford 2005), and we can therefore use this method to investigate the output properties of the individual units stimulated. Compared with the motor fields, which correlate the activity of a unit to the movement vector without making any statement about the downstream connectivity, microstimulation directly addresses the contribution that a particular unit makes toward the generation of the movement. Thus, given the difference between these techniques, we wondered if simulated microstimulation would reveal properties different from both the sensory and motor tuning of the units. In microstimulation experiments, such differences are often attributed to the stimulation-induced, nonphysiological activation of axon pathways (instead of the neurons), but here in our simulations, we were able to constrain activation to only the physiologically relevant outputs.

In this section, we will specify a 0-cm visual desired movement vector to the network (so that the network would naturally not produce any movement), and we keep the visual hand and target input constant, at a zero horizontal and vertical

retinal angle and both at 30-cm distance from the eyes. We then apply simulated microstimulation to individual HLU (second layer) or population output (third) layer units in order to evoke a movement. Microstimulation consisted in setting the specified unit's activity arbitrarily artificially to 2 (we chose a value >1 , the upper limit to a unit activity in ordinary network functioning, in order to ensure large enough stimulation-induced movement vectors). To perform a reference frame analysis, we repeated this procedure for different eye positions so as to allow us to observe the effect of eye position on the simulated microstimulation results.

Four typical results of this analysis are shown in Fig. 9A–D, for microstimulation applied to 4 different HLUs (second) of our example 36-HLU network. Each individual black line represents one movement vector for each eye position, which ranged from -45° to 45° in 5° steps. Intuitively, the starting position of these vectors should change with eye position because the hand position (in shoulder-centered coordinates) has to change to follow eye position. In other words, in order to maintain the same retinal input (same location on the retina) across changes in eye position, hand position in shoulder-centered coordinates must move with the eye. The end points of all stimulation-induced movement vectors were connected by the colored lines. Because our microstimulation changed only the activity of one single unit, this gives an indication as to the nature of the unit's contribution to the movement vector. (Note: Although it is not possible to stimulate a single unit experimentally, the results should in principle be comparable to the effect of activating units with similar properties in a topographically organized region of cortex.)

We observed 4 different types of behaviors. Figure 9A shows an example of HLU #10, where the movement vectors were parallel when microstimulation was applied, irrespective of the change in eye position. We labeled this a fixed-vector movement because the movement vector appears to be in the same direction regardless of different eye positions. (The top part of each panel represents a view from above, the lower part a view from behind.) Another typical result is shown in Figure 9B for microstimulation of HLU #19. In this case, the evoked movement tightly followed eye position, and we, thus, labeled this example gaze dependent. Figure 9C depicts an example for unit #20 that showed an intermediate behavior midway in between the fixed-vector (panel A) and gaze-dependent (panel B) examples. Here, we still observed some rotation of the stimulation-evoked movement vector with eye position but to a smaller extent than eye movement amplitude. Finally, we also found units for which the movement vector evoked through microstimulation converged at a particular location in space (Fig. 9D). We labeled this example a goal-directed movement.

We interpreted the fixed-vector example of Fig. 9A as showing shoulder-centered coordinates because the stimulation-induced movement vector did not depend on eye position but rather produced an approximately constant movement in space. In the typical example in Fig. 9B, the microstimulation-evoked movement vector tightly followed eye position for eye orientation in the horizontal direction. Therefore, this example unit shows a gaze-centered output when tested through simulated microstimulation because the resulting movement vector can be interpreted as being constant in gaze-centered coordinates. As

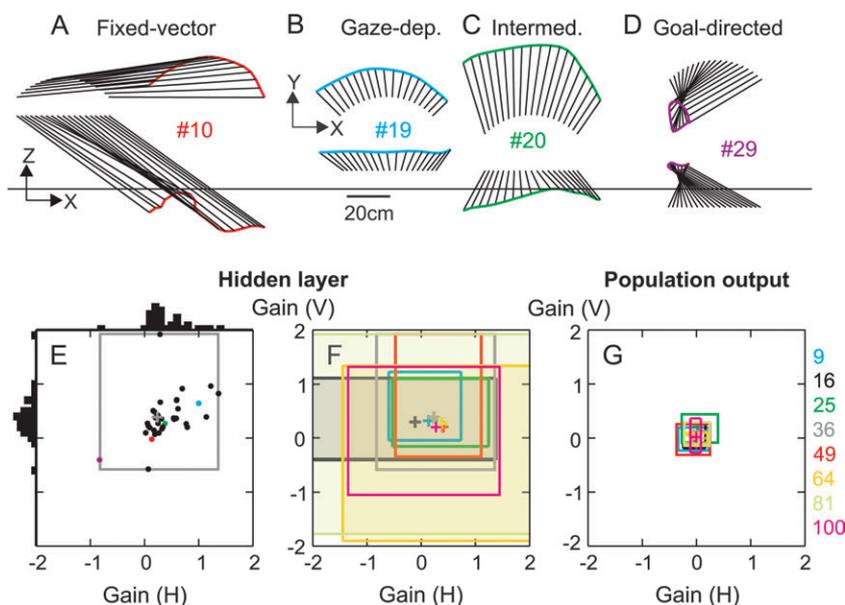


Figure 9. Reference frame analysis through simulated microstimulation of hidden layer and population code units. (A–D) Typical examples of movements generated through simulated microstimulation of individual units under different horizontal eye positions. The movement vector defined in the visual input had zero (0 cm) amplitude. Black lines represent the obtained movement vectors in space (colored bold lines connect endpoints) if the indicated unit (here HLUs of the 36-HLU network) was microstimulated under different horizontal eye positions between -45° and 45° (5° steps), that is, the activity of this unit was artificially set to 2. The top part of the panels represents the top view (X, horizontal; Y, forward) and the lower part shows the behind view (X, horizontal; Z, vertical). We found 4 types of generated movements, that is, fixed-vector movements (panel A), gaze-dependent movements (panel B), a behavior intermediate between fixed-vector and gaze-dependent (panel C) and goal-directed movements (panel D). (E) The horizontal and vertical rotational gain modulation due to horizontal and vertical eye position changes during stimulation, respectively, is shown for each HLU of the 36-HLU network. Colored dots indicate the examples shown in panels (A–D). (F) Summary of the rotational gain ranges for the HLUs of all networks (same representation as Figs 7D,F and 8D,F). (G) The same analysis as in panel (F) but for the population code units.

a consequence, the example in Fig. 9C shows behavior that is intermediate between the predictions a unit working in gaze- and shoulder-fixed coordinates. Finally, Fig. 9D would result from a reference frame that is opposite to gaze centered (one might call it anti-gaze centered) because the rotation of the evoked movement vector is opposite to eye orientation.

To quantify these observations, we proceeded in a similar fashion as for the visual RF and motor field analysis. We performed a regression analysis between the angular deviations of the stimulation-induced movement vectors for different eye positions in the horizontal and vertical directions. This resulted in rotational gain values that could be interpreted with respect to reference frames, that is, a gain of 1 indicates gaze-centered coordinates and a gain of 0 indicates shoulder-centered coordinates. Figure 9E shows the result of this analysis for all HLUs (second) of our example 36-HLU network, plotting the vertical stimulation-induced deviation gain (regression with vertical eye position) as a function of the horizontal gain (regression with horizontal eye position). We observed a large range of different gain values indicating that the hidden layer uses a mixture of different reference frames intermediate between gaze- and shoulder-centered coordinates (see histograms on the axes). This was consistent across all our networks (Fig. 9F). However, the same analysis performed on the population output (third) provided results similar to the example shown in Fig. 9A, and we observed only gain values close to 0 (Fig. 9G). Thus, the output properties of the population output layer only showed close to shoulder-centered coordinates when tested through simulated microstimulation.

Global Encoding Schemes

Up to this point, we have analyzed the different units' apparent reference frames across different electrophysiologically inspired techniques. However, we have only focused on gaze-versus shoulder-centered coordinates and univariate effects (e.g., eye movements in isolation). In this section, we perform 2 additional analyses, first investigating apparent reference frames when including head movements and second addressing relative encoding of movement vectors with respect to eye, hand, and target.

To discriminate between eye-centered, head-centered, and shoulder-centered encoding, we performed a more detailed analysis across all 3 electrophysiological techniques. We used the same analysis techniques as in the previous 3 sections but now changed eye and head orientations. In order to be able to discriminate between eye-, head-, and shoulder-centered encoding, we had to use 3 different conditions, that is, eye-only movements, head-only movements (eye-in-head orientation remained constant), and opposite eye-head movements (like in the VOR). This trivariate analysis explicitly tests for all 3 predictions, whereas one prediction would have to be deduced indirectly from a bivariate analysis, which does not work for the nonlinear behavior of our units (i.e., there is no linear relationship between the three encodings). Table 2 shows the predictions of expected RF shift gains, motor field shift gains, and movement vector rotation gains during microstimulation for the 3 types of movements assuming encoding in either eye-, head-, or shoulder-centered coordinates. We then plotted the model unit gains for changes in eye position, head position, and combined VOR movements in a trivariate plot (each gain corresponding to one axis). In order to best

Table 2

Predicted gains for reference frames

Movements	Ref. frames	Eye	Head	VOR
RFs				
	Eye centered	0	0	0
	Head centered	-1	0	1
	Shoulder centered	-1	-1	0
Motor fields and microstimulation				
	Eye centered	1	1	0
	Head centered	0	1	1
	Shoulder centered	0	0	0

discriminate between the apparent reference frames, we then rotated this 3D plot into a view orthogonal to the plane spanned by the 3 predictions.

Figure 10 shows the result of this analysis in a 3D plot for the HLUs (second layer, panels A-C) and population output (third layer, panels D-F) of the example 36-HLU network and only considering horizontal movements (qualitatively the same results were observed for vertical movements, data not shown). For each technique, the view of this 3D representation was orthogonal to the plane spanned by the 3 predictions of the corresponding technique as shown in Table 2. Colored dots correspond to the 3 predictions and black dots are data points from individual network units. We reproduced our findings from the previous analyses. In addition, we now show that some units reveal close to head-centered behavior; however, most units show encoding in a partially head-centered but closer to shoulder-centered reference frame when probed using motor fields (both for HLUs and population code units) or microstimulation (for HLUs). This was remarkable because there never was any explicit head-centered encoding in the network's input or output. Nevertheless, head-centered-like encoding did emerge in a 3D network involved in visuomotor reference frame transformations for reaching.

Another interesting observation concerned the individual movement gains obtained. Indeed, all 3 movement conditions generally resulted in gain changes. For example, RFs of the population output units also shifted in the head movement condition (Fig. 10D). However, this head movement-related RF shift was along a global direction that was inconsistent with the predictions of head-centered encoding (green dot). Therefore, observing movement-related gain changes does not automatically determine the encoding scheme; rather, a full multivariate analysis is required (see Discussion).

In order to test our modeling approach against known physiological results and to compare our findings with a different view of movement encoding (i.e., absolute vs. relative position encoding), we analyzed the response field gradients. Response field gradients have recently been used to characterize encoding schemes of neurons in the parietal cortex (area MIP, medial intraparietal area) (Buneo et al. 2002) and the PM cortex (Pesaran et al. 2006). Briefly, this method calculates RF changes across eye, hand, and target position and infers the reference frame of a unit from the pattern of modulation observed across the different movements (see also Materials and Methods). It mainly addresses whether a unit encodes the absolute position of one variable (i.e., hand, eye, or target position) alone, independently of any other variable, or whether it uses a relative encoding scheme representing one variable relative to another.

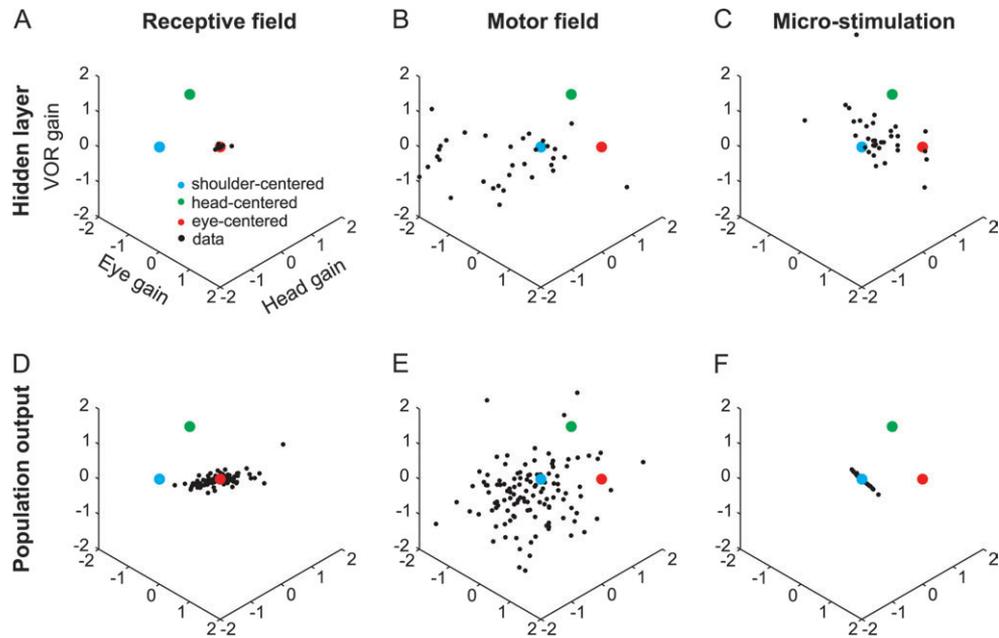


Figure 10. Eye- versus head- versus shoulder-centered frames of reference. The 3D plot depicts gain modulations of each individual unit (black data points) in the example 36-HLU network for eye, head, and VOR movements. Colored data points represent predictions for eye-centered (red), head-centered (green), and shoulder-centered (blue) coordinates. The view of the 3D plot was chosen to be orthogonal to the plane of the 3 predictions. The different panels show results for all 3 electrophysiological techniques, that is, RFs (A, C), motor fields (B, D), and microstimulation (C, F), for HLUs (panels A–C) and population code units (panels D–F).

Figure 11 illustrates the results of the response field gradient analysis for our 100-HLU network (we used the 100-HLU network to have more units and thus better distributions). Panels (A) and (B) show typical HLU (second layer) and population output unit (third layer) activation, respectively, for changes in hand and target positions. For example, the HLU activity in Figure 11A changes with initial hand position but is invariant across target position. This points toward an independent encoding of hand position. Other HLUs also show the inverse pattern, that is, independent encoding of the target position. In contrast, the activity in the typical population output unit (Figure 11B) shows a local maximum for a specific combination of hand and target positions. This was the case for most of the population output units and indicates that these units code the relative position of hand and target and not their individual absolute position, as this was the case in the hidden layer (panel A).

From this pattern of activation across eye, hand, and target positions, we computed the response field gradient for each unit (see Materials and Methods) and analyzed this index with respect to reference frames. The result of this analysis is shown in Figure 11C–D for HLUs (panel C) and population output units (panel D). The first column in Figure 11C depicts the encoding of the target with respect to eye position (the head was fixed in this analysis). As can be seen from the average vector (magenta), HLUs mostly encode target position relative to where the eyes are. The same is true for the encoding of hand position (middle column). However, for the relative encoding of hand and target position (third column), the situation changed. HLUs only showed either absolute hand or absolute target position encoding (like in Fig. 11A), but no relative encoding of hand and target position (as this was the case for the population output example in Fig. 11B). The first 2 columns of Figure 11D show an encoding scheme in the

population output units that was similar to those of the HLUs, that is, hand and target position were encoded relative to the eye. Although, on average, HLUs encoded target position or hand position relative to the eye but not target position relative to the hand, the population output units differ in that they do encode target position relative to the hand (third column of Fig. 11D).

For both a gaze-centered and shoulder-centered reference frame, one expects the same result for the first and second column of panels (C) and (D), that is, an encoding of the target and initial hand position relative to gaze position, which results in a downward population average (see Pesaran et al. 2006). Indeed, regardless of the reference frame used, the position of the hand and target is specified in retinal coordinates, that is, the difference between the spatial hand or target position and current eye (or more generally gaze) position. This was approximately true in our network data. (Note: If initial hand position was encoded relative to the shoulder in the input of our network, the prediction would be different here.) However, for the third column showing the interdependence of hand and target position, the prediction differs between encoding schemes. Although one could argue that relative position codes could exist in any reference frame, it is more likely that these relative codes emerge closer to the motor output. Therefore, we believe that a shoulder-centered representation would be consistent with a relative encoding scheme of hand and target position and predict a downward population average (Pesaran et al. 2006), whereas a gaze-centered encoding would predict independent encoding of hand and target position, that is, a horizontal (pointing to either side) direction for the population (Buneo et al. 2002; Pesaran et al. 2006).

In this view, HLUs (Fig. 11C) show gaze-centered encoding of either the initial hand or target position (third column), but

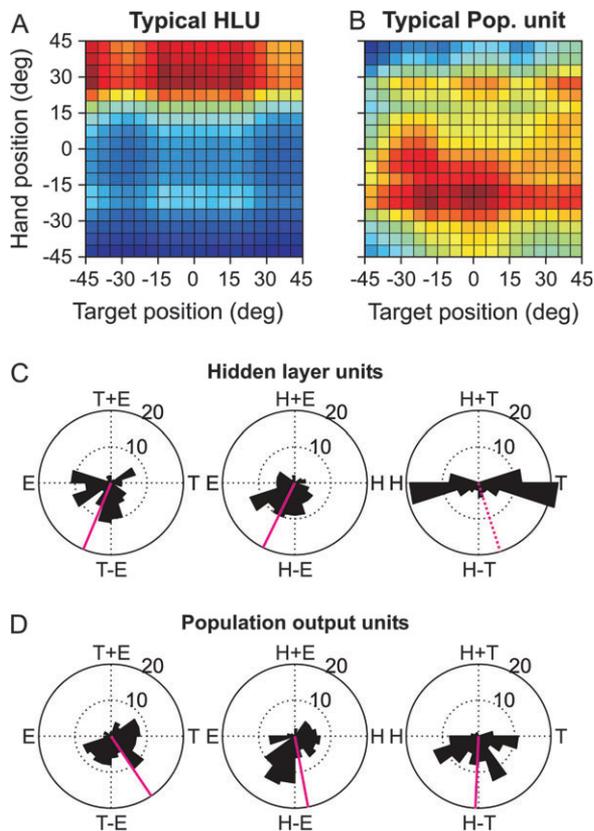


Figure 11. Gradient analysis of reference frames for a 100-HLU network. (A) Activation of a typical HLU for 0° eye position as a function of hand position and target position. Blue indicates not active, red stands for maximally activated. (B) Same representation for a typical population output unit. (C, D) Results of the gradient analysis for HLUs (panel C) and population output (panel D) units. The three columns represent the 3 possible combinations of eye position (E), target position (T), and initial hand position (H). The black rose histogram shows 15° bins. The magenta line represents the population mean over all units.

never both together, whereas population output layer units (Fig. 11D) show on average shoulder-centered coordinates, but with a wide spread toward gaze-centered coordinates (third column). In summary, our network was able to reproduce previous electrophysiological findings from PM and parietal areas, and this even when probing unit properties using relative position codes instead of absolute encoding schemes.

Synthesis of Unit and Layer Properties

One important property of the individual units in our network was the consistent difference in their input–output coding, depending on how they were examined. For example, units in the hidden layer showed purely gaze-centered visual RFs (input coding, Figs 5 and 7E,F), but in their output, they displayed a range of different reference frames distributed between and beyond gaze- and shoulder-centered coordinates. This was true for the motor fields (Fig. 8A–E) and for simulated microstimulation (Fig. 9A–F). Similar observations were made for the population output layer. Thus, different input–output relations were observed within each layer. Interestingly, this was also the case within individual units of each of these layers. In other words, each unit typically displayed different reference frames when tested with different electrophysiological techniques. Thus, each unit performed a fixed input–output transformation (i.e., a fixed mapping between the sensory RF-related inputs

and the resulting output of a given unit as probed by motor fields or microstimulation), and so the visual (input) and motor (output) codes did not align.

These individual transformation modules are essential to the performance of the network. This can be traced to the fundamental aspects of the reference frame transformation that we trained the network to perform. As described in more detail in our previous paper (Blohm and Crawford 2007), the 3D visuomotor transformation for arm movements requires a nonlinear conversion of sensory desired movement vectors in gaze-centered coordinates into nonidentical movement vectors in shoulder-centered coordinates, as a function of eye and head configuration. A network composed of fixed, but aligned local input–output mappings (for each individual unit), could not perform such a transformation no matter how these local transformations were combined. Mathematically, this means that if the sensory and motor reference frames are equal, then no reference frame transformation had taken place. Thus, the transformation modules are used as components for the overall transformation at the population level. Only by combining fixed transformation modules in different ways, could one achieve different global transformations at the population level.

The remaining question is how were these fixed transformation modules correctly combined to produce the global reference frame transformation? The answer to this question relies in the second important property in our network, which was the existence of gain modulation. Indeed, the activity of many of the HLUs (second layer) and population output (third) layer units in our neural network was largely modulated by eye, head, and hand position signals in a gain-like fashion (Andersen et al. 1985; Zipser and Andersen 1988; Salinas and Abbott 2001). This has been observed in real neurons in the brain (e.g., in PPC) and indicates the potential involvement of this part of the brain in reference frame transformations for reaching (Andersen et al. 1985; Galletti et al. 1995; Battaglia-Mayer et al. 2001; Buneo et al. 2002). Gain modulation was present throughout the network, across all layers (hidden layer: Figs 5F and 8E; population output: Fig. 8G) and was mediated by the additional inputs to the network (eye position, head position, vergence) which when combined with the fixed input–output mappings produced the gain modulation. We hypothesized that these gain modulations allowed the network to weight the contribution of each unit’s fixed transformation (different input–output coding) in a way to accurately produce the complete overall transformation.

Figure 12 tests and illustrates this hypothesis by directly showing how gain modulation in the hidden (second) and population output (third) layer contributes to the construction of the overall reference frame transformation. We chose 2 typical situations, 1) using the same retinal movement vector to produce 3 different movements in space for different head roll angles (Fig. 12A) and 2) producing the same movement vector in space from 3 different visual inputs and head roll angles (Fig. 12B). Because in the first condition (Fig. 12A), all inputs but the head roll angle were held constant, potential modulations of HLU (second layer) activity must be due to a change in head roll. Therefore, head roll weighted the contribution of these HLUs fixed input–output transformations to the overall reference frame conversion. However, the activity modulations in the population output (third) layer units (Fig. 12A) resulted from a combination of gain modulation and different motor outputs. Conversely, in the second condition (Fig. 12B), the

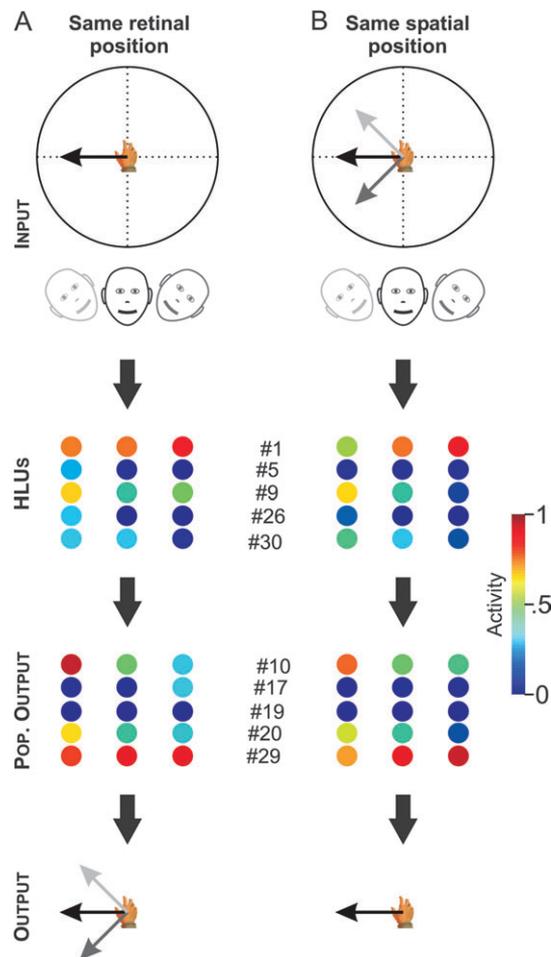


Figure 12. Activity gain modulations in the network. (A) The same retinal target position produces 3 different movement vectors for 3 different head roll angles. The activity of 5 typical HLUs (same units as in Fig. 5) and population output units (same units as in Figs 8 and 9) is shown for each condition. Eye position was zero and initial hand position was at the fovea. Target and hand position were on a 50-cm distant tangential plane. (B) Three different retinal target positions are presented for 3 different head roll angles so as to produce the same movement vector in space. The activation for the same units as in panel (A) is shown.

motor output was constant and modulations of the population output (third) layer units' activity therefore had to originate from gain modulation through the input. In this case, the third layer units received gain modulation to weight the contribution of those units to the motor output, thus weighting the fixed transformations modules differently. Finally, the activity modulations in the HLUs (second layer) (Fig. 12B) resulted from a combination of gain modulation and different visual inputs.

As can be observed from Figure 12A, most of the example HLUs' (second layer) activity was—at least to some extent—modulated by the head roll angle (these are the same example units as shown in Fig. 4), whereas all other inputs to the network were held constant. The activity of most HLUs (second layer) changed for different head roll angles and as we have shown in Figures 5 and 7E–F, this change in activity was not due to a RF shift but could only result from gain modulation. As already mentioned before and shown in Figure 8G, we also observed gain modulation in the population output layer. This can be seen in Figure 12B, where the typical example population output units' activity was modulated by head position (and also the visual target position) despite

producing the same movement vector. The same principle applies for any arbitrary combination of visual target, hand position, and eye-head configuration. Thus, from our analysis, we conclude that gain modulation at every level of the network was crucial to combine the contributions of different fixed transformation modules of individual units as parallel components in order to produce the global transformation required in the network's overall input–output mapping.

To summarize these results at the population level of our network (Fig. 2D), units in the hidden (second) and population output (third) layer had visual RFs with activations that—at the population level—covered all different parts of visual space. In addition, the HLUs had purely gaze-centered inputs (visual RFs) and their output properties reflected mixed, intermediate properties between gaze- and shoulder-centered coordinates (motor fields and microstimulation). The population output units also showed mixed, intermediate input properties. When tested using motor fields, the output properties of this layer showed mixed reference frames; however, simulated microstimulation displayed purely shoulder-centered output coordinates for the population output layer. Therefore, we conclude that individual units transformed information through fixed input–output relationships. The complete reference frame transformation was performed by summation across these fixed transformation modules at each processing level, weighted in a gain-like fashion using eye, head, vergence, and hand position signals.

Discussion

We trained a physiologically inspired artificial neural network to perform the 3D visuomotor transformation from gaze-centered inputs to a shoulder-centered output as required for geometrically accurate reach planning (Fig. 2A, Blohm and Crawford 2007). The network was able to perform this complex nonlinear transformation and it did so in a gradual, distributed fashion. Different methodologies, that is, whether we used visual RFs (testing the input properties of units), motor fields, or microstimulation (both testing different output properties of units) provided different reference frames within the same units in a particular network layer (Fig. 2D). These results help to highlight and explain the fundamental difference between the 3 main techniques available to systems electrophysiologists and demonstrate that it is critical to be aware of these differences when comparing results obtained with different approaches. In addition, separately probing the sensory and motor reference frames of individual units allowed us to show how individual units implement fixed reference frame transformation modules that can be combined in a gain-weighted fashion to produce an overall transformation at the population level. Some of these observations have been made before either theoretically or experimentally; this is the first study to synthesize all of them and show how they arise naturally as a consequence of solving the specific geometric problems for 3D reach.

Neurophysiological Significance

Our working hypothesis was that the hidden layer would represent PPC areas that have been suggested to be involved in transforming these early visual signals into motor plans (e.g., Caminiti et al. 1998; Burnod et al. 1999; Snyder 2000; Buneo

et al. 2002; Battaglia-Mayer et al. 2003; Crawford et al. 2004; Buneo and Andersen 2006). On the other hand, we hypothesized that the population output layer represented the PM cortex (Kalaska et al. 1997; Kakei et al. 2001, 2003; Scott 2001), and we designed this layer—through the readout mechanism—to mimic neural properties in the PM cortex (see Materials and Methods).

Within the limits of this simple network architecture, we were able to reproduce and explain many findings of real neurons in the frontal-parietal network and clarify certain controversies in the field. For example, the visual RFs of different regions in the superior parietal lobe are known to encode target position in gaze-centered coordinates (Johnson et al. 1996; Batista et al. 1999; Caminiti et al. 1999; Battaglia-Mayer et al. 2001, 2003; Buneo et al. 2002; Pesaran et al. 2006). The same scheme is likely used to encode hand position (Crammond and Kalaska 2000; Buneo et al. 2002). However, aligning the neural activity in those same areas with the hand movement vector (motor fields) revealed mixed, intermediate properties that fall between reference frames (Battaglia-Mayer et al. 2001, 2003). In addition, electrical microstimulation of PPC produces complex movements that might reflect such “hybrid” reference frames (Cooke et al. 2003; Stepniewska et al. 2005). Our model explains these seemingly incompatible observations between the sensory and motor reference frames of units in the same PPC areas as being the result of the inherent properties of individual neurons involved in reference frame transformations. Therefore, using different experimental techniques addressing distinct inherent properties of a neural network can lead to incompatibilities of observed results.

One interesting aspect of our model was that some units of our network displayed an apparent head-centered reference frame when probed using motor fields or microstimulation (Fig. 10). This was remarkable because neither in the input nor in the output there was information explicitly using this reference frame. Furthermore, although head-movement-related gain modulation has been described for neurons in PPC (Brochier et al. 1995, 2003), head-centered reference frames have only been reported when probing the units’ motor fields (Battaglia-Mayer et al. 2003) and have never been observed when mapping out visual RFs (e.g., Batista et al. 1999), which is exactly what our model predicted. These consistencies between physiological observations and our model predictions underline the relevance of our model for interpreting neural response properties.

This is more than just a methodological point because—as shown here—understanding the difference of results obtained across techniques is crucial for identifying neurons directly involved in reference frame transformations: the spatial input and output properties of these neurons should not match. In general, when trying to identify a site involved in reference frame transformations, we should look for an area whose units show 1) input properties that are relatively weighted toward the sensory frame, 2) output properties that are relatively weighted toward the shoulder frame, and 3) gain modulations related to the relative orientation and location between these 2 frames.

Where can one observe such properties in the existing literature? Eye position gain fields occur throughout the parietofrontal network for reach (Andersen et al. 1985; Snyder 2000; Battaglia-Mayer et al. 2001; Buneo and Andersen 2006). However, our model may help to understand the different and seemingly contradictory properties seen in some parts of this

system, like PM. It is largely believed that PM encodes movements in extrinsic coordinates (Kalaska et al. 1997; Kakei et al. 2001, 2003; Scott 2001). This interpretation is based on the motor tuning properties (motor fields) of the neurons in PM and is consistent with a motor plan for the arm in shoulder-centered coordinates (Boussaoud and Wise 1993; Shen and Alexander 1997). In these studies, eye and head position were kept constant, and therefore, no eye/head position effects were reported for the motor field properties of PM units, consistent with our model. However, seemingly contradictory to the view of extrinsic coordinates in PM is the finding that the visual RFs in PM are modulated by eye position (Boussaoud et al. 1993; Pesaran et al. 2006). We show here that this is an inherent property of units in a population that is actively involved in a visuomotor transformation network. This is also consistent with current electrophysiological recordings showing eye position modulations of neural activity in PM (Mushiake et al. 1997; Boussaoud et al. 1998; Joffrais and Boussaoud 1999; Cisek and Kalaska 2002). Most interestingly, different reference frames between gaze- and shoulder-centered coordinates have recently been reported in PM neurons during visually guided reaching (Batista et al. 2007), just as predicted by our network. In an additional analysis (Fig. 11), we were also able to reproduce recent findings concerning the differences between neurons in PPC and PM with respect to reference frames (Pesaran et al. 2006), where PM units displayed relative position codes (i.e., hand position relative to target position, consistent with shoulder-centered coordinates), whereas PPC showed absolute position coding (i.e., of either hand or target position, consistent with gaze-centered coordinates). Through analogy with our model, this directly implicates PM in the visuomotor transformations for 3D reach.

Similar observations arise when comparing visual input properties and the motor output evoked during electrical microstimulation. For example, the transformations for large head-free gaze shifts must deal with a nonlinear transformation analogous to the one studied here. Gaze-related units in the supplementary eye fields (SEFs) primarily show eye-centered visual RFs (Russo and Bruce 1996), but microstimulation of the SEF revealed coding schemes in multiple effector-based frames ranging from eye centered to head centered to body centered (Schlag and Schlag-Rey 1987; Martinez-Trujillo et al. 2004; Park et al. 2006). This closely agrees with the behavior observed here in the intermediate and output layers of our network, except for reach rather than gaze.

When PM is stimulated, complex movements that do not fit a shoulder-centered reference frame have been evoked (Graziano et al. 2002). However, the long stimulation trains used in that study might have resulted in the activation of more dynamic properties within the PM and motor network that are beyond motor planning but rather address motor execution and control (Churchland and Shenoy 2007). Again, as these areas generally show a simpler eye-centered organization in their visual input and more complex goal-directed behavior in their output, the current theoretical framework tends to implicate them in transformations from sensory to motor space.

General Methodological Implications

Sensory tuning, motor tuning, and microstimulation do not reveal the same information about individual units, that is, we have demonstrated that a same unit can have a different sensory-, motor-, and stimulation-induced reference frame. This

means that results obtained using different electrophysiological techniques should be compared with care. For example, in networks involved in reference frame conversions, these 3 techniques should not provide the same results, not just because of experimental noise (although this will always be a confounding factor in neurophysiological experiments) but because these differences are fundamental to the underlying mechanism. If the sensory and motor reference frames were the same, then no reference frame transformation would take place. One thus expects to find different reference frames within the same area in the brain when testing with different techniques. This is crucial, on the one hand, for avoiding misinterpretation or apparent contradictions between these techniques and, on the other hand, for capitalizing on the full potential of such comparisons.

This observation does not only hold true for reference frames. In general, if a neuron does indeed participate in a purposeful way in a certain computation, then different input and output properties must be expected; otherwise, this neuron would merely transmit information but not transform it in any way, that is, no computation would be performed. Therefore, finding different apparent reference frames when using different probing techniques is not contradictory at all. Rather, it provides a useful tool to electrophysiologists when searching for the neural substrates underlying complex sensorimotor functions. Of course, finding the same input and output properties does not prove that a particular neuron is not involved in a computation; such conclusions could only be drawn from a population analysis.

One important observation was that when probing reference frames, one has to perform a multivariate analysis. This is particularly true when one needs to discriminate between different reference frames in reaching networks. For example, our analysis in Figure 10 has shown that a unit that displays tuning shifts with head movements does not necessarily encode information in a head-centered reference frame. This observation was striking for RF shifts in the population output layer (Fig. 10D). There was a large range of RF shift gain changes related to head movements; however, multivariate analysis using different combinations of eye and head movements revealed that there was no head-centered representation present here. Therefore, we stress that in order to perform discriminations between different potential reference frames (as long as more than 2 possibilities exist in the system), a multivariate analysis has to be performed. Failure to do so would lead to wrong conclusions. For example, if we had only analyzed head movements in the attempt to discriminate between shoulder-centered and head-centered reference frames in the population output layer of our network (Fig. 10D), we would obtain mixed reference frames, intermediate between head- and shoulder-centered rather than intermediate between eye and shoulder centered. This is the case because the head carries the eyes, and thus, only moving the head will not distinguish between eye- and head-centered reference frames; hence, the need for a multivariate analysis.

We observed different reference frames when computing motor fields or simulating microstimulation. For example, the population output layer encoded movement vectors in mixed reference frames when probed with motor fields rather than the shoulder-centered coordinates one might have expected. In our network, this was the case because there was enough redundancy so that when the contributions of the population

layer neurons were summed, the resulting movement was still shoulder centered. In contrast, when simulating microstimulation of an individual population output unit, the unit was not being driven in its normal fashion by its hidden layer inputs and responses appear shoulder-centered because the unit is tied directly to the output. The question then arises, what does the difference in the results obtained from recording motor fields and from microstimulation mean? Indeed, both techniques address the output of a unit and one might expect to find at least similar if not the same results. However, there are important conceptual differences in both techniques. First, motor fields look at the natural contribution of a unit's activity to the motor output, whereas microstimulation artificially produces a motor action. Second, motor fields do not make any statement about the magnitude of the contribution of a unit to the motor action because the output is produced by the whole network. In contrast, microstimulation probes specifically how a unit (or a group of neurons in the brain) contributes to action, leaving the rest of the upstream network relatively unaffected. And third, microstimulation allows inducing unit activities that are not achieved naturally by the network (in particular in the real brain where a group of neurons is activated together), whereas this is not the case when simply recording natural activity and aligning it with the motor output, as done to compute motor fields.

The reference frame analysis we used is, in some ways, a very limited tool to characterize the activity of neurons in the brain or in our network, and therefore, the usefulness of the approach could be questioned. Unfortunately, we do not have a better tool available to date. Describing the apparent frame of reference of a unit is a particularly valuable technique for quantitatively describing the response properties of neurons involved in sensorimotor transformations. The different input-output relationships then tell us something about the units' functions and, as such, provide useful information. However, this does not mean that individual units in a network actually perform an explicit reference frame transformation. Perhaps, more importantly, irrespective of the actual meaning of "intermediate" reference frames, we provide qualitative and quantitative predictions as to what to look for and what to expect in the real brain when trying to identify areas involved in reference frame transformations. Therefore, we believe that our approach to analyze the apparent reference frames of a unit has practical applications.

Comparison to Previous and Alternative Models

The use of a complete 3D geometric approach in our model is much more than a trivial extension of previous work using 1D or 2D geometry because the incorporation of the real nonlinear 3D geometry of this system brings new principles into play that are not present in linear approximations. For example, we did not observe a strict alignment of retinal and extraretinal contributions in the RF shifts of our population output units as was the case in previous studies (Salinas and Abbott 1995). Rather, we found that a purely horizontal eye movement could elicit vertical RF shifts as required to compensate for a tilted visual input resulting from torsional positions of the eyes and/or head. Such predictions are not mistakes or oddities, they are crucial for solving the nonlinear problems mentioned above. The most notable difference between ours and previous approaches is that a reach model based on 2D geometry only

uses translations, whereas the 3D geometry has to deal with the multiplicative complexity of noncommutative, nonlinear rotations (Pouget and Sejnowski 1997; Tweed et al. 1999), for example, when both the eyes and head move while their centers of rotation are not aligned. Furthermore, the 3D geometry is not only much more realistic (and actually the only correct approach) but is also particularly important for computing the distance component of the reach; however, this analysis is beyond the scope of the present paper and will be described elsewhere.

In addition, we show that a multivariate analysis is needed in order to perform correct reference frame discriminations (see previous section). Another novel observation was that different methodologies (e.g., RFs, motor fields, and microstimulation) can give rise to different findings regarding reference frame for the same unit (see previous section). Finally, we developed a new way of analyzing neural networks that was only possible when using a realistic 3D model.

Despite the new properties that emerge in 3D, our network reproduces many findings that have been reported in previous algebraic modeling and neural network studies, for example, a gain-like modulation of unit activity with eye, head, and hand position (Zipser and Andersen 1988; Brotchie et al. 1995, 2003; Salinas and Abbott 1995, 2001; Snyder et al. 1997; Xing and Andersen 2000), purely gaze-centered RFs in the HLUs (Zipser and Andersen 1988; Xing and Andersen 2000), “intermediate” reference frames (Xing and Andersen 2000)—although these were only observed when combining inputs encoded in different reference frames—or shifting RFs in the motor output layer (Salinas and Abbott 1995). This is despite the fact that we did not use “basis function networks,” as previously done (e.g., Zipser and Andersen 1988; Salinas and Abbott 1995; Pouget and Sejnowski 1997), that are known to be able to approximate any nonlinear function (e.g., Buhmann 2003). Whereas in basis function networks one designs the input sensitivity of a unit, for example, by specifying how it should respond to targets presented with respect to a certain portion of space, we left our network free to come up with its own solution. In doing so, we avoided combinatorial explosion of the number of HLUs required as this is the case when using radial basis functions (Pouget and Snyder 2000); instead, we left it up to the network to develop the best possible set of “basis-like” functions. The most noticeable difference in the results obtained was that RFs never shifted in our network; shifting RFs have, however, been reported when using radial basis functions (Xing and Andersen 2000), a property that is—however—inconsistent with PPC data (e.g., Buneo and Andersen 2006). Another observation that has previously been described was the misalignment of sensory and motor coordinates, but this has only been done for the much simpler saccadic eye movement system (Smith and Crawford 2005). Thus, for the most part, our study does not negate these previous findings but rather reinforces them by extending them to 3D, and more importantly, showing why they must exist in a system that must deal with the real-life problems of acting in 3D.

For example, whereas the importance of gain fields is sometimes claimed to be ambiguous for linear 2D approximations (Colby and Goldberg 1999), they become essential to solve the complex orientation-dependent sensory-motor relations seen in 3D space. Indeed, the formal need of gain fields depends on the way neurons encode position signals. If the brain used a vector coding scheme, gain fields would not be

needed for 2D visuomotor conversions because linear, additive operations are sufficient to obtain remapping, updating, or reference frame translations (Pouget and Sejnowski 1997; Colby and Goldberg 1999). This consideration has led to doubt as to whether gain modulations serve reference frame transformations or rather reflect eye/head position-dependent attentional modulations in areas such as PPC (Colby and Goldberg 1999). In nonlinear 3D geometry, however, the vector coding scheme gets into trouble, and we show here that now the use of gain fields is crucial in order to achieve reference frame transformations. It is the nonlinear aspect that requires the use of gain fields, as is also the case for computations that use basis functions (Salinas and Abbott 1995; Pouget and Sejnowski 1997).

Perhaps the most general theoretical contribution of our study for the understanding of reach is to show how these various principles must work in unison to solve the real-world problems of moving in 3D space. Using the complete 3D geometry and a new detailed reference frame analysis provides a unifying framework on how reference frame transformations can be achieved through fixed transformation modules and weighted, distributed processing.

Our model is far from complete. Many extensions of our neural network model are possible and necessary in the future. Computing with noisy neurons has been shown to add valuable insight into the potential neural mechanisms of the visuomotor transformation (Salinas and Abbott 1995; Pouget and Snyder 2000; Deneve et al. 2001). It would also be useful to test more biologically plausible learning rules and to add proprioceptive input of the hand position to the model to see how spatial inputs in different modes are integrated in a neural network. Adding proprioceptive hand position should not change our overall findings though because it has been shown that PPC always encodes hand position in visual coordinates, regardless of whether the hand is seen or not (Buneo et al. 2002), which leads to the conclusion that the proprioceptive-to-visual coordinate transformation may occur before PPC, maybe in S1 (primary somatosensory area). In addition, adding movement kinetics to the output coding would allow us to address the transformation of the extrinsic movement plan into intrinsic muscle activations.

Further Predictions

Using a 3D geometric approach, we attempted to design a neural network that was physiologically realistic in its input-output architecture. Based on the relative simplicity of the emerging features and the many parallels between our findings and those in the experimental literature, we propose that the mechanisms that we have observed are not just theoretical constructs but in fact are fundamental to how the real brain plans reaches.

In order to test this hypothesis, more experiments are required. One of the most important would be to provide direct comparisons between sensory tuning, motor tuning, and microstimulation results for a single unit/brain site. We predict that units involved in a reference frame transformation process will display different reference frames in their input (RF) and output (motor field, microstimulation) properties. These differences should be associated with the presence of gain fields (Andersen et al. 1985; Zipser and Andersen 1988; Salinas and Abbott 1995). This is true not only for areas involved in the

visuomotor transformation for reaching. We believe that our main conclusions are general enough to apply to all reference frame transformations in the brain (e.g., eye-to-head or multi-segment transformations along the joints of the arm).

Second, as mentioned earlier, we expect microstimulation and motor fields to display “different” output properties when tested with both techniques. The congruence between microstimulation and motor field results might thus give us indications about the specificity of the contribution of a unit to a certain task. In contrast, differences might be indicative of complex computations where many different parameters can influence the motor field, whereas microstimulation (at least at high currents) would remain relatively unaffected.

Third, we suggest that microstimulation of PPC units during feed-forward motor planning should affect the movement in different reference frames. This is also consistent with transcranial magnetic stimulation and brain lesion studies that show a variety of reference frames involved in different aspects of the visuomotor transformation for reaching (Khan et al. 2005, 2007; van Donkelaar and Adams 2005; Vesia et al. 2006).

Our neural network model might also be useful to investigate how specific brain lesions affect the sensorimotor transformation. We suggest that virtual lesions induced in such artificial neural networks could provide valuable insight into the origin of the deficits of parietal damage patients. For example, lesions to HLUs with unilateral visual RFs might produce movement deficits similar to those observed in optic ataxia patients (Khan et al. 2005, 2007). However, experimental evidence is needed to specifically address the deficits due to an internally “damaged” geometrical model for reaching under different eye-head positions. As previously suggested (Blohm and Crawford 2007), this model could be used to provide valuable insights into other visuomotor transformation deficits related to damage of several parts of the brain. Pathologies of interest include strabismus, cerebellar damage, motor learning disorders, Alzheimer’s (or other neurodegenerative diseases), and vestibular system damage.

The present artificial neural network study is thus much more than an abstract theoretical investigation of how reference frame transformations could occur in distributed processing networks within our brain; rather, we believe it to be fundamental to better understand the current neurophysiological investigations and brain-damaged patients. However, it is only the first step in modeling the 3D sensorimotor geometry of reach. Next steps would be to extend this model to include multisensory (visual, auditory, somatosensory) inputs of target and hand position, and/or a more complete model of limb kinematics and dynamics at the output end, simulated using the real 3D geometry of these systems. This should provide additional insights into the role of neural networks in multisensory integration and the inverse kinematics/dynamics of motor control.

Supplementary Material

Supplementary Methods can be found at: <http://www.cercor.oxfordjournals.org/>.

Funding

The Canadian Institutes of Health Research (CIHR); Marie Curie International fellowship within the Sixth European Community Framework Program and CIHR, Canada (to G.B.).

Notes

We thank Dr A. Z. Khan for helpful comments on the manuscript. J.D.C. holds a Canada Research Chair. *Conflict of Interest*: None declared.

Address correspondence to Dr J. Douglas Crawford, Centre for Vision Research, York University, 4700 Keele Street, Toronto, Ontario M5P 2L3, Canada. Email: jdc@yorku.ca.

References

- Andersen RA, Essick GK, Siegel RM. 1985. Encoding of spatial location by posterior parietal neurons. *Science*. 230:456–458.
- Baraduc P, Guigon E, Burnod Y. 2001. Recoding arm position to learn visuomotor transformations. *Cereb Cortex*. 11:906–917.
- Batista AP, Buneo CA, Snyder LH, Andersen RA. 1999. Reach plans in eye-centered coordinates. *Science*. 285:257–260.
- Batista AP, Santhanam G, Yu BM, Ryu SI, Afshar A, Shenoy KV. 2007. Reference frames for reach planning in macaque dorsal premotor cortex. *J Neurophysiol*. 98(2):966–83.
- Battaglia-Mayer A, Caminiti R, Lacquaniti F, Zago M. 2003. Multiple levels of representation of reaching in the parieto-frontal network. *Cereb Cortex*. 13:1009–1022.
- Battaglia-Mayer A, Ferraina S, Genovesio A, Marconi B, Squatrito S, Molinari M, Lacquaniti F, Caminiti R. 2001. Eye-hand coordination during reaching. II. An analysis of the relationships between visuomanual signals in parietal cortex and parieto-frontal association projections. *Cereb Cortex*. 11:528–544.
- Blohm G, Crawford JD. 2007. Computations for geometrically accurate visually guided reaching in 3-D space. *J Vis*. 7:1–22.
- Bockisch CJ, Haslwanter T. 2001. Three-dimensional eye position during static roll and pitch in humans. *Vision Res*. 41:2127–2137.
- Boussaoud D, Barth TM, Wise SP. 1993. Effects of gaze on apparent visual responses of frontal cortex neurons. *Exp Brain Res*. 93:423–434.
- Boussaoud D, Joffrais C, Bremmer F. 1998. Eye position effects on the neuronal activity of dorsal premotor cortex in the macaque monkey. *J Neurophysiol*. 80:1132–1150.
- Boussaoud D, Wise SP. 1993. Primate frontal cortex: effects of stimulus and movement. *Exp Brain Res*. 95:28–40.
- Brotchie PR, Andersen RA, Snyder LH, Goodman SJ. 1995. Head position signals used by parietal neurons to encode locations of visual stimuli. *Nature*. 375:232–235.
- Brotchie PR, Lee MB, Chen DY, Lourens M, Jackson G, Bradley WG, Jr. 2003. Head position modulates activity in the human parietal eye fields. *Neuroimage*. 18:178–184.
- Buhmann MD. 2003. Radial basis functions: theory and implementations. Cambridge (UK): Cambridge University Press.
- Buneo CA, Andersen RA. 2006. The posterior parietal cortex: sensorimotor interface for the planning and online control of visually guided movements. *Neuropsychologia*. 44(13):2594–606.
- Buneo CA, Jarvis MR, Batista AP, Andersen RA. 2002. Direct visuomotor transformations for reaching. *Nature*. 416:632–636.
- Burgess N. 2006. Spatial memory: how egocentric and allocentric combine. *Trends Cogn Sci*. 10:551–557.
- Burnod Y, Baraduc P, Battaglia-Mayer A, Guigon E, Koechlin E, Ferraina S, Lacquaniti F, Caminiti R. 1999. Parieto-frontal coding of reaching: an integrated framework. *Exp Brain Res*. 129:325–346.
- Caminiti R, Ferraina S, Mayer AB. 1998. Visuomotor transformations: early cortical mechanisms of reaching. *Curr Opin Neurobiol*. 8:753–761.
- Caminiti R, Genovesio A, Marconi B, Mayer AB, Onorati P, Ferraina S, Mitsuda T, Giannetti S, Squatrito S, Maioli MG, et al. 1999. Early coding of reaching: frontal and parietal association connections of parieto-occipital cortex. *Eur J Neurosci*. 11:3339–3345.
- Churchland MM, Shenoy KV. 2007. Temporal complexity and heterogeneity of single-neuron activity in premotor and motor cortex. *J Neurophysiol*. 97:4235–4257.
- Cisek P, Kalaska JF. 2002. Modest gaze-related discharge modulation in monkey dorsal premotor cortex during a reaching task performed with free fixation. *J Neurophysiol*. 88:1064–1072.

- Cohen YE, Andersen RA. 2002. A common reference frame for movement plans in the posterior parietal cortex. *Nat Rev Neurosci*. 3:553-562.
- Colby CL, Goldberg ME. 1999. Space and attention in parietal cortex. *Annu Rev Neurosci*. 22:319-349.
- Cooke DF, Taylor CS, Moore T, Graziano MS. 2003. Complex movements evoked by microstimulation of the ventral intraparietal area. *Proc Natl Acad Sci USA*. 100:6163-6168.
- Crammond DJ, Kalaska JF. 2000. Prior information in motor and premotor cortex: activity during the delay period and effect on pre-movement activity. *J Neurophysiol*. 84:986-1005.
- Crawford JD. 1994. The oculomotor neural integrator uses a behavior-related coordinate system. *J Neurosci*. 14:6911-6923.
- Crawford JD, Cadera W, Vilis T. 1991. Generation of torsional and vertical eye position signals by the interstitial nucleus of Cajal. *Science*. 252:1551-1553.
- Crawford JD, Guitton D. 1997. Visual-motor transformations required for accurate and kinematically correct saccades. *J Neurophysiol*. 78:1447-1467.
- Crawford JD, Henriques DY, Vilis T. 2000. Curvature of visual space under vertical eye rotation: implications for spatial vision and visuomotor control. *J Neurosci*. 20:2360-2368.
- Crawford JD, Medendorp WP, Marotta JJ. 2004. Spatial transformations for eye-hand coordination. *J Neurophysiol*. 92:10-19.
- Crawford JD, Vilis T. 1992. Symmetry of oculomotor burst neuron coordinates about Listing's plane. *J Neurophysiol*. 68:432-448.
- Deneve S, Latham PE, Pouget A. 2001. Efficient computation and cue integration with noisy population codes. *Nat Neurosci*. 4:826-831.
- Flash T, Sejnowski TJ. 2001. Computational approaches to motor control. *Curr Opin Neurobiol*. 11:655-662.
- Fu QG, Suarez JI, Ebner TJ. 1993. Neuronal specification of direction and distance during reaching movements in the superior precentral premotor area and primary motor cortex of monkeys. *J Neurophysiol*. 70:2097-2116.
- Fukushima K, Harada C, Fukushima J, Suzuki Y. 1990. Spatial properties of vertical eye movement-related neurons in the region of the interstitial nucleus of Cajal in awake cats. *Exp Brain Res*. 79:25-42.
- Fukushima K, Ohashi T, Fukushima J, Kase M. 1992. Ocular torsion produced by unilateral chemical inactivation of the interstitial nucleus of Cajal in chronically labyrinthectomized cats. *Neurosci Res*. 13:301-305.
- Galletti C, Battaglini PP, Fattori P. 1995. Eye position influence on the parieto-occipital area PO (V6) of the macaque monkey. *Eur J Neurosci*. 7:2486-2501.
- Goldberg ME, Bruce CJ. 1990. Primate frontal eye fields. III. Maintenance of a spatially accurate saccade signal. *J Neurophysiol*. 64:489-508.
- Graziano MS, Taylor CS, Moore T. 2002. Complex movements evoked by microstimulation of precentral cortex. *Neuron*. 34:841-851.
- Haslwanter T, Straumann D, Hess BJ, Henn V. 1992. Static roll and pitch in the monkey: shift and rotation of Listing's plane. *Vision Res*. 32:1341-1348.
- Henriques DY, Crawford JD. 2002. Role of eye, head, and shoulder geometry in the planning of accurate arm movements. *J Neurophysiol*. 87:1677-1685.
- Henriques DY, Klier EM, Smith MA, Lowy D, Crawford JD. 1998. Gaze-centered remapping of remembered visual space in an open-loop pointing task. *J Neurosci*. 18:1583-1594.
- Henriques DY, Medendorp WP, Gielen CC, Crawford JD. 2003. Geometric computations underlying eye-hand coordination: orientations of the two eyes and the head. *Exp Brain Res*. 152:70-78.
- Hepp K. 1995. Theoretical explanations of Listing's law and their implication for binocular vision. *Vision Res*. 35:3237-3241.
- Johnson PB, Ferraina S, Bianchi L, Caminiti R. 1996. Cortical networks for visual reaching: physiological and anatomical organization of frontal and parietal lobe arm regions. *Cereb Cortex*. 6:102-119.
- Jouffrais C, Boussaoud D. 1999. Neuronal activity related to eye-hand coordination in the primate premotor cortex. *Exp Brain Res*. 128:205-209.
- Jurgens R, Becker W, Kornhuber HH. 1981. Natural and drug-induced variations of velocity and duration of human saccadic eye movements: evidence for a control of the neural pulse generator by local feedback. *Biol Cybern*. 39:87-96.
- Kakei S, Hoffman DS, Strick PL. 2001. Direction of action is represented in the ventral premotor cortex. *Nat Neurosci*. 4:1020-1025.
- Kakei S, Hoffman DS, Strick PL. 2003. Sensorimotor transformations in cortical motor areas. *Neurosci Res*. 46:1-10.
- Kalaska JF, Crammond DJ. 1992. Cerebral cortical mechanisms of reaching movements. *Science*. 255:1517-1523.
- Kalaska JF, Scott SH, Cisek P, Sergio LE. 1997. Cortical control of reaching movements. *Curr Opin Neurobiol*. 7:849-859.
- Keith GP, Smith MA, Crawford JD. 2007. Functional organization within a neural network trained to update target representations across 3-D saccades. *J Comput Neurosci*. 22(2):191-209.
- Khan AZ, Crawford JD, Blohm G, Urquizar C, Rosetti Y, Pisella L. Forthcoming 2007. Influence of initial hand and target position on reach errors in optic ataxic and normal subjects. *J Vis*. 7(5):8.1-16.
- Khan AZ, Pisella L, Vighetto A, Cotton F, Luaute J, Boisson D, Salemme R, Crawford JD, Rossetti Y. 2005. Optic ataxia errors depend on remapped, not viewed, target location. *Nat Neurosci*. 8:418-420.
- Khokhotva M, Ono H, Mapp AP. 2005. The cyclopean eye is relevant for predicting visual direction. *Vision Res*. 45:2339-2345.
- King WM, Fuchs AF, Magnin M. 1981. Vertical eye movement-related responses of neurons in midbrain near interstitial nucleus of Cajal. *J Neurophysiol*. 46:549-562.
- Klier EM, Crawford JD. 1998. Human oculomotor system accounts for 3-D eye orientation in the visual-motor transformation for saccades. *J Neurophysiol*. 80:2274-2294.
- Lehky SR, Sejnowski TJ. 1990. Neural network model of visual cortex for determining surface curvature from images of shaded surfaces. *Proc R Soc Lond B Biol Sci*. 240:251-278.
- Marsaglia G. 1972. Choosing a point from the surface of a sphere. *Ann Math Stat*. 43:645-646.
- Martinez-Trujillo JC, Medendorp WP, Wang H, Crawford JD. 2004. Frames of reference for eye-head gaze commands in primate supplementary eye fields. *Neuron*. 44:1057-1066.
- Mascaro M, Battaglia-Mayer A, Nasi L, Amit DJ, Caminiti R. 2003. The eye and the hand: neural mechanisms and network models for oculomanual coordination in parietal cortex. *Cereb Cortex*. 13:1276-1286.
- Medendorp WP, Crawford JD. 2002. Visuospatial updating of reaching targets in near and far space. *Neuroreport*. 13:633-636.
- Merriam EP, Colby CL. 2005. Active vision in parietal and extrastriate cortex. *Neuroscientist*. 11:484-493.
- Messier J, Kalaska JF. 2000. Covariation of primate dorsal premotor cell activity with direction and amplitude during a memorized-delay reaching task. *J Neurophysiol*. 84:152-165.
- Muller ME. 1959. A note on a method for generating points uniformly on N-dimensional spheres. *Comm Assoc Comput Mech*. 2:19-20.
- Mushiake H, Tanatsugu Y, Tanji J. 1997. Neuronal activity in the ventral part of premotor cortex during target-reach movement is modulated by direction of gaze. *J Neurophysiol*. 78:567-571.
- Naka KI, Rushton WA. 1966a. An attempt to analyse colour reception by electrophysiology. *J Physiol*. 185:556-586.
- Naka KI, Rushton WA. 1966b. S-potentials from colour units in the retina of fish (Cyprinidae). *J Physiol*. 185:536-555.
- Naka KI, Rushton WA. 1966c. S-potentials from luminosity units in the retina of fish (Cyprinidae). *J Physiol*. 185:587-599.
- Nikara T, Bishop PO, Pettigrew JD. 1968. Analysis of retinal correspondence by studying receptive fields of binocular single units in cat striate cortex. *Exp Brain Res*. 6:353-372.
- Ohzawa I, DeAngelis GC, Freeman RD. 1997. Encoding of binocular disparity by complex cells in the cat's visual cortex. *J Neurophysiol*. 77:2879-2909.
- Ono H, Barbeito R. 1982. The cyclopean eye vs. the sighting-dominant eye as the center of visual direction. *Percept Psychophys*. 32:201-210.
- Ono H, Mapp AP, Howard IP. 2002. The cyclopean eye in vision: the new and old data continue to hit you right between the eyes. *Vision Res*. 42:1307-1324.
- Pare M, Wurtz RH. 1997. Monkey posterior parietal cortex neurons antidromically activated from superior colliculus. *J Neurophysiol*. 78:3493-3497.

- Park J, Schlag-Rey M, Schlag J. 2006. Frames of reference for saccadic command tested by saccade collision in the supplementary eye field. *J Neurophysiol.* 95:159-170.
- Pellionisz A, Llinas R. 1985. Tensor network theory of the metaorganization of functional geometries in the central nervous system. *Neuroscience.* 16:245-273.
- Pellionisz AJ, Ramos CF. 1993. Geometrical approach to neural net control of movements and posture. *Prog Brain Res.* 97:245-256.
- Pesaran B, Nelson MJ, Andersen RA. 2006. Dorsal premotor neurons encode the relative position of the hand, eye, and goal during reach planning. *Neuron.* 51:125-134.
- Pettigrew JD, Nikara T, Bishop PO. 1968. Binocular interaction on single units in cat striate cortex: simultaneous stimulation by single moving slit with receptive fields in correspondence. *Exp Brain Res.* 6:391-410.
- Poggio GE. 1995. Mechanisms of stereopsis in monkey visual cortex. *Cereb Cortex.* 5:193-204.
- Poggio GF, Fischer B. 1977. Binocular interaction and depth sensitivity in striate and prestriate cortex of behaving rhesus monkey. *J Neurophysiol.* 40:1392-1405.
- Pouget A, Sejnowski TJ. 1994. A neural model of the cortical representation of egocentric distance. *Cereb Cortex.* 4:314-329.
- Pouget A, Sejnowski TJ. 1997. Spatial transformations in the parietal cortex using basis functions. *J Cogn Neurosci.* 9:222-237.
- Pouget A, Snyder LH. 2000. Computational approaches to sensorimotor transformations. *Nat Neurosci.* 3(Suppl):1192-1198.
- Ren L, Khan AZ, Blohm G, Henriques DY, Sergio LE, Crawford JD. 2006. Proprioceptive guidance of saccades in eye-hand coordination. *J Neurophysiol.* 96:1464-1477.
- Riedmiller M, Braun H. 1993. A direct adaptive method for faster backpropagation learning: the RPROP algorithm. *IEEE International Conference on Neural Networks.* San Francisco (CA): IEEE pages 586-591; Meeting Date: 03/28/1993 - 04/01/1993.
- Rushworth MF, Taylor PC. 2006. TMS in the parietal cortex: updating representations for attention and action. *Neuropsychologia.* 44:2700-2716.
- Russo GS, Bruce CJ. 1996. Neurons in the supplementary eye field of rhesus monkeys code visual targets and saccadic eye movements in an oculocentric coordinate system. *J Neurophysiol.* 76:825-848.
- Salinas E, Abbott LF. 1994. Vector reconstruction from firing rates. *J Comp Neurosci.* 1:89-107.
- Salinas E, Abbott LF. 1995. Transfer of coded information from sensory to motor networks. *J Neurosci.* 15:6461-6474.
- Salinas E, Abbott LF. 1996. A model of multiplicative neural responses in parietal cortex. *Proc Natl Acad Sci USA.* 93:11956-11961.
- Salinas E, Abbott LF. 2001. Coordinate transformations in the visual system: how to generate gain fields and what to compute with them. *Prog Brain Res.* 130:175-190.
- Schlag J, Schlag-Rey M. 1987. Evidence for a supplementary eye field. *J Neurophysiol.* 57:179-200.
- Scott SH. 2001. Vision to action: new insights from a flip of the wrist. *Nat Neurosci.* 4:969-970.
- Scott SH. 2003. The role of primary motor cortex in goal-directed movements: insights from neurophysiological studies on non-human primates. *Curr Opin Neurobiol.* 13:671-677.
- Scott SH, Gribble PL, Graham KM, Cabel DW. 2001. Dissociation between hand motion and population vectors from neural activity in motor cortex. *Nature.* 413:161-165.
- Shen L, Alexander GE. 1997. Preferential representation of instructed target location versus limb trajectory in dorsal premotor area. *J Neurophysiol.* 77:1195-1212.
- Smith MA, Crawford JD. 2001. Self-organizing task modules and explicit coordinate systems in a neural network model for 3-D saccades. *J Comput Neurosci.* 10:127-150.
- Smith MA, Crawford JD. 2005. Distributed population mechanism for the 3-D oculomotor reference frame transformation. *J Neurophysiol.* 93:1742-1761.
- Snyder LH. 2000. Coordinate transformations for eye and arm movements in the brain. *Curr Opin Neurobiol.* 10:747-754.
- Snyder LH, Batista AP, Andersen RA. 1997. Coding of intention in the posterior parietal cortex. *Nature.* 386:167-170.
- Sober SJ, Sabes PN. 2003. Multisensory integration during motor planning. *J Neurosci.* 23:6982-6992.
- Sober SJ, Sabes PN. 2005. Flexible strategies for sensory integration during motor planning. *Nat Neurosci.* 8:490-497.
- Soechting JF, Tillery SIH, Flanders M. 1991. Transformation from head-to shoulder-centered representation of target direction in arm movements. *J Cogn Neurosci.* 2:32-43.
- Somani RA, DeSouza JF, Tweed D, Vilis T. 1998. Visual test of Listing's law during vergence. *Vision Res.* 38:911-923.
- Stepniewska I, Fang PC, Kaas JH. 2005. Microstimulation reveals specialized subregions for different complex movements in posterior parietal cortex of prosimian galagos. *Proc Natl Acad Sci USA.* 102:4878-4883.
- Suzuki Y, Buttner-Ennever JA, Straumann D, Hepp K, Hess BJ, Henn V. 1995. Deficits in torsional and vertical rapid eye movements and shift of Listing's plane after uni- and bilateral lesions of the rostral interstitial nucleus of the medial longitudinal fasciculus. *Exp Brain Res.* 106:215-232.
- Tehovnik EJ, Slocum WM, Schiller PH. 2003. Saccadic eye movements evoked by microstimulation of striate cortex. *Eur J Neurosci.* 17:870-878.
- Todorov E. 2000. Direct cortical control of muscle activation in voluntary arm movements: a model. *Nat Neurosci.* 3:391-398.
- Todorov E. 2002. Cosine tuning minimizes motor errors. *Neural Comput.* 14:1233-1260.
- Todorov E, Jordan MI. 2002. Optimal feedback control as a theory of motor coordination. *Nat Neurosci.* 5:1226-1235.
- Tsutsui K, Taira M, Sakata H. 2005. Neural mechanisms of three-dimensional vision. *Neurosci Res.* 51:221-229.
- Tweed D. 1997. Visual-motor optimization in binocular control. *Vision Res.* 37:1939-1951.
- Tweed DB, Haslwanter TP, Happe V, Fetter M. 1999. Non-commutativity in the brain. *Nature.* 399:261-263.
- van Donkelaar P, Adams J. 2005. Gaze-dependent deviation in pointing induced by transcranial magnetic stimulation over the human posterior parietal cortex. *J Mot Behav.* 37:157-163.
- Van Rijn LJ, Van den Berg AV. 1993. Binocular eye orientation during fixations: Listing's law extended to include eye vergence. *Vision Res.* 33:691-708.
- Vesia M, Monteon JA, Sergio LE, Crawford JD. 2006. Hemispheric asymmetry in memory-guided pointing during single-pulse transcranial magnetic stimulation of human parietal cortex. *J Neurophysiol.* 96:3016-3027.
- Xing J, Andersen RA. 2000. Models of the posterior parietal cortex which perform multimodal integration and represent space in several coordinate frames. *J Cogn Neurosci.* 12:601-614.
- Zipser D, Andersen RA. 1988. A back-propagation programmed network that simulates response properties of a subset of posterior parietal neurons. *Nature.* 331:679-684.