Check for updates

Models of the Posterior Parietal Cortex Which Perform Multimodal Integration and Represent Space in Several Coordinate Frames

Jing Xing and Richard A. Andersen
California Institute of Technology

Downloaded from http://direct.mit.edu/jocn/article-pdf/12/4/601/1758964/089892900562363.pdf by guest on 18 May 2021

Abstract

■ Many neurons in the posterior-parietal cortex (PPC) have saccadic responses to visual and auditory targets. The responses are modulated by eye position and head position. These findings suggest that PPC integrates multisensory inputs and may provide information about saccadic targets represented in different coordinate frames. In addition to an eye-centered output representation, PPC may also project to brain areas which contain head-centered and body-centered representations of the space. In this report, possible coordinate transformations in PPC were examined by comparing several sets of models of PPC, each having different representations in the output layer: (i) an eye-centered map only; (ii) a head-centered map only; (iii) an eye-centered map and a head-centered map; and (iv) an eye-centered map, a head-centered map, and a body-centered map. These output maps correctly encoded saccades to visual and auditory targets through training. The units in the hidden layers of the models exhibited the following properties: (1) The units had gain fields (GFs) for eye position, and also for head position if the model had a body-centered output representation; (2) As the result of the GF and the nonlinear activation function of the units, the hidden layers often employed "intermediate" coding, e.g., the hidden units coded targets partially in eye-centered coordinates and, partially, in head-centered coordinates; (3) Different types of coordinate transformations in these models were carried out by different relationships between the receptive fields (RFs) and the GFs of the hidden units; and (4) The properties of PPC neurons are in better accordance with the hidden units of the models that had multiple-output representations than the models that had only one single-output representation. In conclusion, the results show that the GF is an effective mechanism for performing coordinate transformations. The models also suggest that neurons with intermediate coding are to be expected in the process of coordinate transformations. ■

INTRODUCTION

Physiological recordings and lesion studies have demonstrated that the posterior-parietal cortex (PPC) is involved in coordinate transformations for representing spatial information (Andersen, Essick, & Siegel, 1987; Mountcastle, Lynch, Georgopoulos, Sakata, & Acuna, 1975). By combining sensory information with extraretinal information, such as signals of eye position or head position, the brain can form abstract representations of the space interposed between the sensory input and the motor output (Zipser & Andersen 1988; Andersen et al., 1985). Increasingly experimental evidence demonstrates that PPC is very important in such information combinations. Neurons in PPC respond to saccades to visual or auditory targets (Mazzoni, Bracewell, Barash, & Andersen, 1996). Moreover, responses of PPC neurons to visual or auditory targets are modulated by eye position and head position, and such modulations are important for coordinate transformations (Brotchie, Andersen, Snyder, & Goodman, 1995; Stricanne et al., 1996; Synder, Brotchie, & Andersen, 1993). For instance, an eye-centered representation can be formed by subtracting information about the eye position from the location of an auditory target encoded in head-centered coordinates; a head-centered representation can be formed by adding information about the eye position and the visual target location on the retinas; and a body-centered representation can be formed by adding information about the eye position, the head position, as well as the retinal target location. Furthermore, experimental data suggest that more than one representation of the space may exist in the parietal cortex (Mazzoni & Andersen, 1995; Synder et al., 1993). Psychological experiments found that different coordinate references were used for different motor tasks (Helms Tillery, Flanders, & Soechting, 1991; Soechting, Flanders, & Helms Tillery, 1990; Soechting & Flanders, 1989a, b). Given that PPC neurons receive information about eye position and head position, and possibly body position, PPC may provide the motor system with multiple-repre-

© 2000 Massachusetts Institute of Technology

sentations of space. It is interesting to determine what kinds of circuits are required for various coordinate transformations.

A number of computational models have been proposed to account for coordinate transformations. Zipser and Andersen (1988) developed a network model, which combined visual signals and eye position signals onto a head-centered output layer. They found that visual signals and eye position signals interacted to form "gain fields (GFs)" in which the amplitudes of visual responses were modulated by the eye position. The GFs of the hidden units in the model were similar to those found in neurons of area 7a, a subregion of PPC. Krommenhoek, Van Opstal, Gielen and Gisbergen (1993) trained a three-layered network model to code motor errors of visual targets topographically in the output layer. They found that the hidden units encoded motor errors in a distributed fashion. A model developed by Xing, Stricanne and Andersen (1994) performed visual-auditory integrations and remapped visual and auditory target locations onto a motor-error representation for saccades at the output layer. The hidden units in this model were tuned to both target location and motor error, similar to the neurons reported in PPC (Stricanne, Andersen, & Mazzoni, 1996). All these models were made on the assumption that the cortex provides only one type of representation at the output level. Pouget and Sejnowski (1995) used basis functions to model PPC neurons. They demonstrated that a set of basis functions could provide both eye-centered and head-centered output representations. To understand the role of PPC in sensorimotor integration, we explored several possible models of coordinate transformations, which might occur in PPC. Simulations in this report were aimed at the following questions: (1) Can PPC provide multiple representations of space? (2) Are there any common principles involved in various transformations? (3) What are the intrinsic differences between neurons involved in different transformations?

We modeled sensorimotor integrations in PPC at several levels: (1) Neurons receive visual or auditory inputs and provide a single-output representation of space; (2) Neurons receive both visual and auditory inputs and provide a single-output representation; and (3) Neurons receive bimodal inputs and provide multiple-representations in different reference frames. Although these may not include all the possible integrations in PPC, the analysis of a subset of models at various integration levels may highlight the intrinsic mechanisms of coordinate transformations in PPC. Parts of this report have been presented in abstract form (Xing et al., 1994).

Description of the Models

Structure of the Models

Since several models were examined in this report, we first introduce the common structures for all the models, then describe each model in detail. Each model consists of three abstract layers: the 2-D input maps, the middle layer (i.e., the hidden layer), and the 2-D output maps. Physiologically, these layers correspond to inputs to PPC, PPC, and the motor/premotor areas to which PPC projects. Every unit in the input layer is connected to all the hidden units, and every hidden unit is connected to all the output units. The connection weights are initially set to small random values and are adjusted through training.

Input layer. The input layer contains a visual map coding visual targets in eye-centered coordinates, an auditory map coding acoustic targets in head-centered coordinates, and eye position units, as well as head position units, which linearly encode eye and head positions. An $80^\circ \times 80^\circ$ visual input map was modeled using an 8×8 array of units. Receptive fields of these units were represented by a Gaussian distribution with a 1/e width of 15°. The centers of the receptive fields (RFs) were equally spaced over the 8×8 grid with 10° spacing. The auditory map was modeled in the same way as the visual map, except the auditory units code auditory targets in head-centered coordinates. These input units may correspond to a subset of sensory neurons in PPC (Mazzoni et al., 1996).

Most eye position neurons in PPC respond monotonically to horizontal and vertical eye positions in head-centered coordinates. This feature was modeled using four sets of eight units. Two sets of these units linearly encoded the horizontal component of eye position with positive and negative slopes. The other two sets encoded the vertical component in a similar manner. The slope and intercept for each unit's response were chosen randomly. Eye positions were, thus, encoded linearly in the activity of these units. Like eye position, head position also modulates the responses of many PPC neurons (Brotchie et al., 1995). Thus, we used four sets of eight head position units to encode head position in a similar manner as coding the eye position except head position was encoded in body-centered coordinates.

Hidden layer. Twenty hidden layer units were typically used in the simulations presented in this report unless specified otherwise. The hidden units received activities from input units of all input channels. The output activation of the hidden units was calculated by first summing all inputs and then calculating the output as a sigmoidal function of the total input.

The activation in an output unit is:

$$\text{Output} = 1/[1 + \exp(-\text{net})]$$

where net = sum of weighted inputs + bias.

The sigmoid was chosen as an output function, because it resembles the operation performed by

Downloaded from http://direct.mit.edu/jocn/article-pdf/12/4/601/1759359/089892900562363.pdf by guest on 18 May 2021

actual neurons, which sum inputs, have a threshold, and saturate with high levels of activity. The sigmoid function limits output activity to a range between 0 and 1. In the middle region of its dynamic range, the sigmoid approximates a linear function.

Output layer. The output layer contains several representations: an eye-centered map, a head-centered map, and a body-centered map. The eye-centered and head-centered maps are similar to the corresponding input maps, with an 8×8 array of output units to represent output signals topographically for each map. The body-centered map was also modeled in a similar manner with 8×8 units. Each of the output maps covers an $80^\circ \times 80^\circ$ space. The activation of output units, like the hidden units, was governed by the sigmoid function.

We use the following symbols in model descriptions:

V: retinal location of visual targets;
A: head-centered location of auditory targets;
E: position of the eyes in head-centered coordinates;
H: position of the head in body-centered coordinates;
ME: motor error in eye-centered coordinates (output representation appropriate for eye movement):
HE: head-position error in head-centered coordinates (output representation appropriate for head movement):
BE: body-position error in body-centered coordinates (output representation appropriate for body movement)

For a given visual target V, the output maps are specified with ME = V for the eye-centered map, HE = V + E for the head-centered map, and BE = V + E + H for the body-centered map. For an auditory target location A, the output maps are specified with ME = A − E, HE = A, and BE = A + H.

Models of Four Different Levels of Integration in PPC

We approached our goals by exploring models at different integration levels. All the models have access to eye position in the input layer. For the first three levels, only the eye-centered and the head-centered output maps were considered. Level-1 models are for unimodal neurons. Different sets of hidden units receive inputs of either visual or auditory target locations and project to one of the two output maps, respectively (Figure 1A). Level-2 models use bimodal inputs. Different sets of hidden units receive both visual and auditory inputs and project to one of the two output maps (Figure 1B). The Level-3 model has bimodal inputs and multiple-output representations. All the hidden units project to both output maps simultaneously (Figure 1C). The Level-4 model has additional access to head-position information, and its output layer contains an eye-centered map, a head-centered map and a body-centered map.

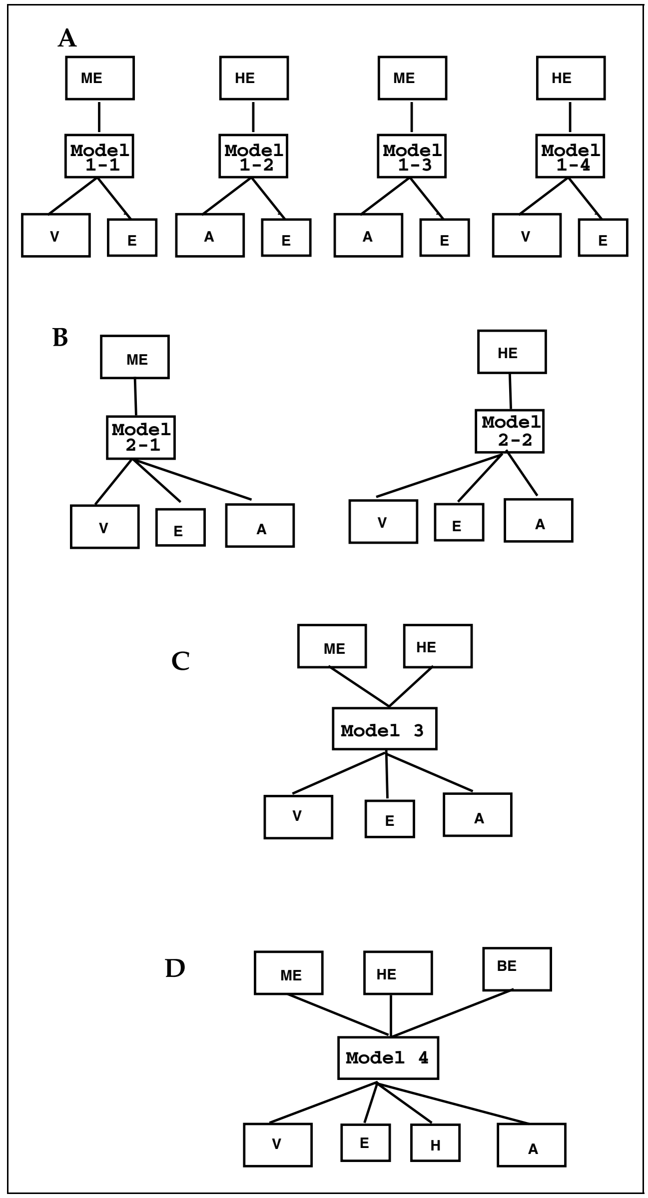


Figure 1. The diagrams of the models. Each model is composed of three layers: 2-D input maps, 2-D output maps, and a hidden layer. Input maps are: V—Visual inputs, A—Auditory inputs, E—Eye position, and H—Head position. Output maps are: ME—saccadic motor error map (eye-centered), HE—head-centered map, and BE—body-centered map. Each input unit is connected to every unit in the hidden layer, each hidden unit projects to every output unit. A, B, C, D show the models of the four different levels of integration.

Models of Level-1. There are four models at this level, as shown in Figure 1A. Model 1-1 remaps visual inputs to eye-centered outputs: ME = V; Model 1-2 remaps auditory inputs to the head-centered representation: HE = A; Model 1-3 remaps auditory inputs to the eye-

Downloaded from ... on 18 May 2021

centered outputs: ME = A − E; Model 1-4 remaps visual inputs to the head-centered representation: HE = V + E. Model 1-1 and Model 1-2 simply remap input signals onto the output layer in the same coordinate frame; while Model 1-3 and Model 1-4 require coordinate transformations.

Models of Level-2. The assumption for Level-2 models is that neurons in PPC are bimodal, i.e., the neurons receive both auditory and visual inputs. The neurons are assumed to segregate into two sets: one set projects to a ME map and the other to a HE map. Correspondingly, two models are needed at this level, as shown in Figure 1B. Model 2-1 receives auditory and visual inputs and remaps them to the eye-centered representation: ME = V and ME = A − E. Model 2-2 remaps visual and auditory inputs to the head-centered representations: HE = V + E and HE = A.

The model of Level-3. The assumption is that individual PPC neurons integrate visual and auditory inputs and project to both eye-centered and head-centered output representations. Therefore, as shown in Figure 1C, one model at this level is sufficient for all the transformations in Level-1 and Level-2. Model 3 has access to auditory and visual inputs. The two output maps of Model 3 are: ME = V, ME = A − E, and HE = A, HE = V + E.

The model of Level-4. In addition to eye position, head position also modulates the responses of many PPC neurons (Brotchie et al., 1995). Head-position information is necessary for a body-centered representation. This is simulated in Model 4. Compared to Model 3, Model 4 receives additional head-position signals (H) and has one more output map, the body-centered representation: BE = V + E + H, BE = A + H. Figure 1D shows the diagram of Model 4.

Training Method and Learning Algorithm

The process of training consists of repeatedly presenting input patterns to the network, comparing the actual output with the desired output, and adjusting synaptic weights accordingly. For each training cycle, the location of either a visual or an auditory target, and an eye position in the orbit or head position were picked at random and were converted into the formats used for the inputs to the model. These inputs were propagated to the output units. The error pattern was calculated as the difference between the

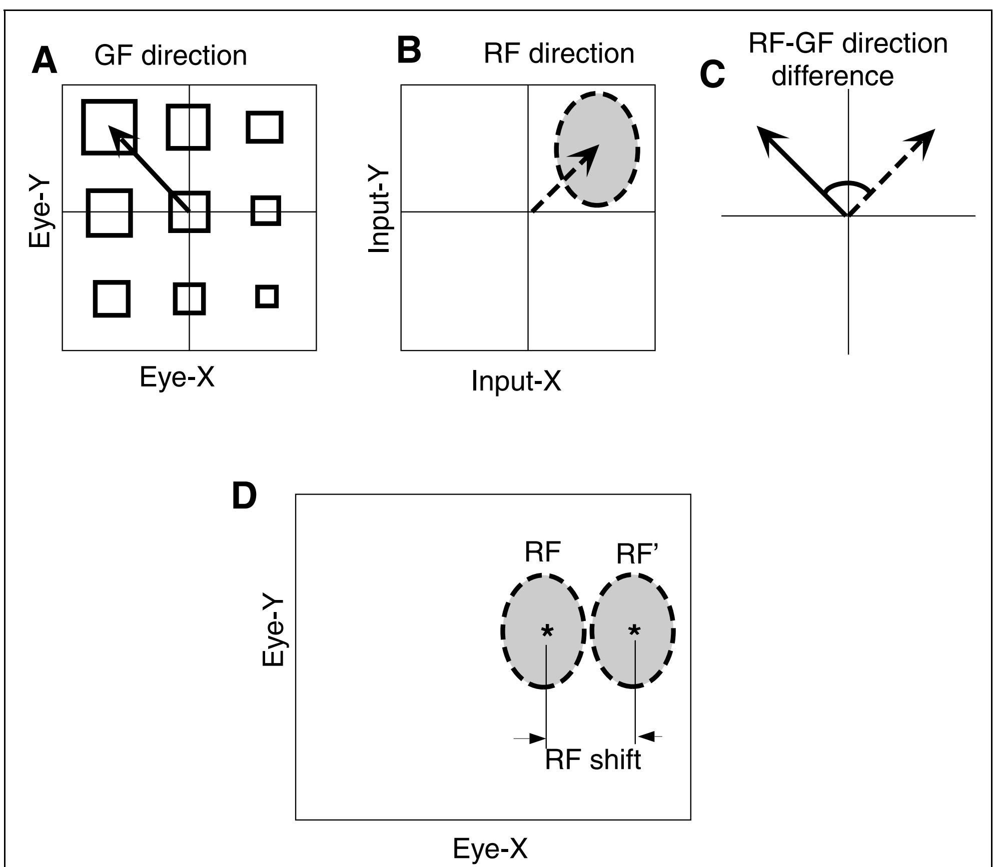
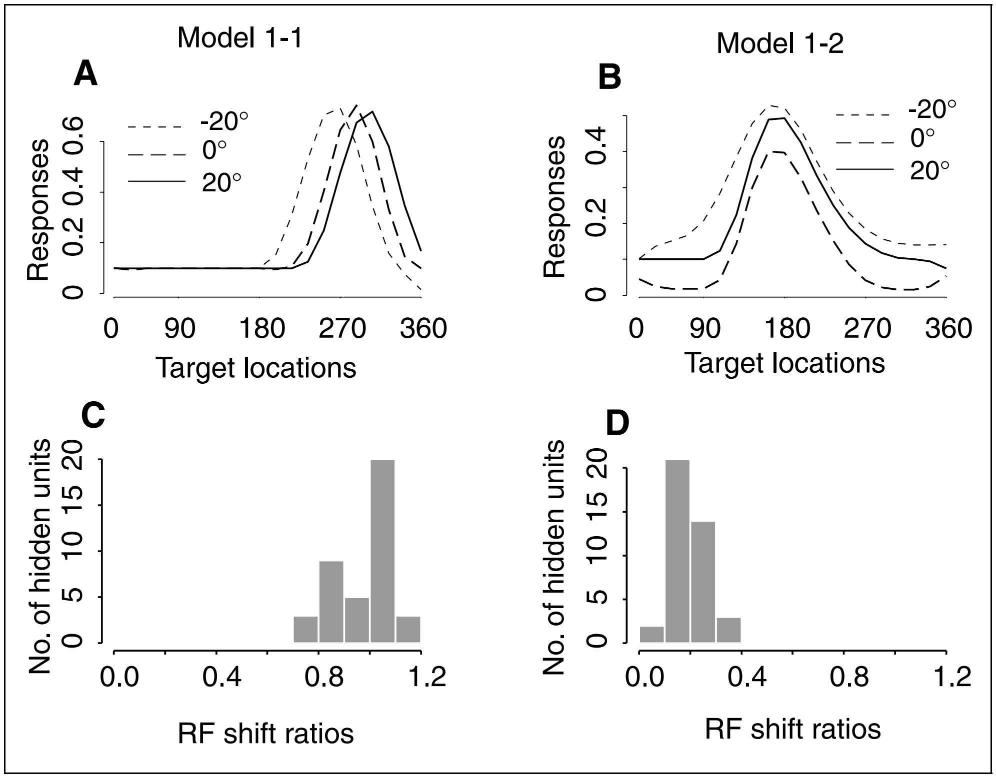


Figure 2. The measurements of RF–GF direction difference and RF shift. (A) The illustration of a GF direction. The size of the squares indicates the responses at different eye positions. The GF direction is from the center of fixation to the eye position of maximal response. (B) The illustration of a RF direction as indicated by the arrow. The RF is indicated with the dashed area and the RF direction is from the center of the input map to the center of the RF. (C) The RF–GF direction difference. (D) RF shift, computed as the horizontal distance of the RF centers in head-centered coordinates for horizontal eye-position shifts. The same calculation is made for the vertical RF shift.

Downloaded from http://direct.mit.edu/jocn/article-pdf/12/4/601/1759323/089892900562372.pdf by guest on 18 May 2021

Figure 3. Receptive fields shift with eye position. (A, B) 1-D RF curves of a typical hidden unit in (A) Model 1-1 and (B) Model 1-2, measured as the unit's responses to targets presented at different angular positions of the input map. Different curves were obtained at three eye positions (−20°, 0°; 0°, 0°; 20°, 0°). (C, D) Histograms of RF shift ratios for all the hidden units of Model 1-1 (C) and Model 1-2 (D). Vertical and horizontal shift ratios were computed for each unit and plotted in the same histogram.



Downloaded from http://direct.mit.edu/jocn/article-pdf/12/4/601/1759559/089892900562363.pdf by guest on 18 May 2021

actual output pattern and the desired output pattern. This error pattern was then used to modify the connection weights of the model using the backpropagation learning algorithm (Rumelhart, Hinton, & Williams, 1986). Training was repeated until two conditions were satisfied: (1) The mean square error between the actual output pattern and the expected output pattern could not be decreased with further training; and (2) A performance error, defined as the difference between the center of mass of the expected output and the network output, is less than 4° for any given combination of inputs within the modeled space. Since an allowed performance error of 4° was typically used for training monkeys to make saccades, the second condition was to assure that a trained model had comparable performance with behaving monkeys. It also assured that all the models after training have approximately the same performance level.

Measurements

After the training was completed, the response properties of the hidden units were examined and compared across different models. The following are the typical measurements made on the hidden units.

Gain field (GF). The response amplitude of a unit to a target is modulated by eye position. The GF of a unit is the 2-D plot of the unit's response to a target presented in the unit's RF measured against the different eye positions. Figure 2A illustrates a GF. The size of the squares is proportional to the response measured at the corresponding eye position.

Receptive field (RF). Most trained units responded selectively to targets presented in a local area of the input maps, as illustrated with the shaded area in Figure 2B. We defined the RF of a hidden unit as the input area within which a target could evoke a response greater than 50% of the unit's maximal response. Unless otherwise specified, RFs were measured when the eye position and the head position pointed to the center of fixation, i.e., E = 0 and H = 0.

GF–RF relationship. The direction of a RF was calculated as the vector direction from the center of the input map to the center of the RF, as illustrated in Figure 2B. The center of a RF was computed as "the center of mass" of the unit's 2-D response across the input map. In the cases when a RF center coincides with the center of the input map, the unit is labeled as "nondirectional". The direction of a GF was calculated as the best-tuned direction of the GF relative to the central eye position, as illustrated in Figure 2A. The GF–RF alignment of a unit was, thus, defined as the direction difference of the GF and the RF of a unit, as shown in Figure 2C. The direction difference varies between 0 and 180°.

RF shift ratio. One of the main goals of this report is to examine the coordinate frames in which PPC neurons encode saccades. Ideally, if hidden units encode targets in eye-centered frames, their RFs would shift with changes in eye position; if neurons encode targets in head-centered frames, then the RFs should not change with eye position. The shift of RFs can, thus, be used as a coordinate index. For a given hidden unit, the RF shift was computed as the distance between the centers of mass of the RF profiles measured at different eye positions. This is illustrated in Figure 2D. Since the eye position was encoded in head-centered coordinates, the RF shift was also computed in head-centered coordinates. The RF shift ratio was then computed as the shift of RF centers in head-centered coordinates divided by the corresponding change in the eye position. For each unit, we calculated RF shift ratios along the horizontal and vertical axis, respectively.

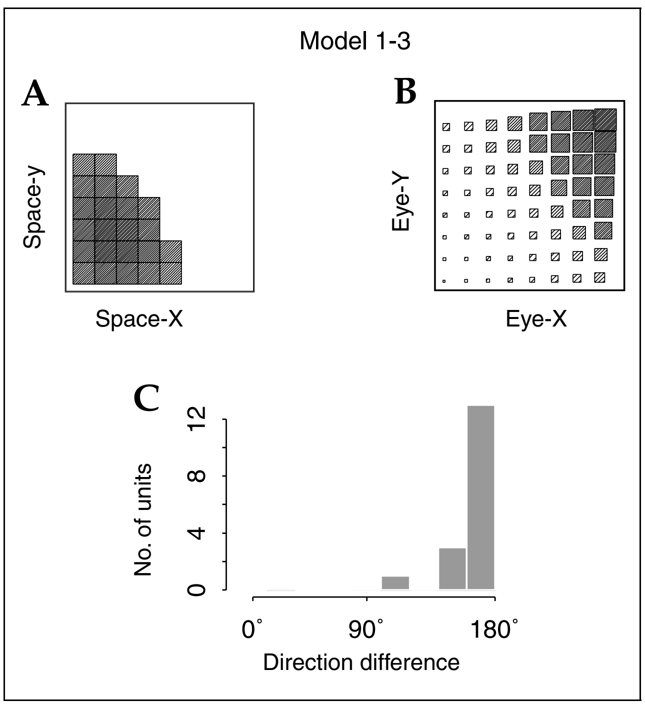


Figure 4. The receptive fields and gain fields of the hidden units in Model 1-3. (A) The receptive field of a typical hidden unit in Model 1-3. The gray level of the small squares is proportional to the evoked responses of the hidden unit. (B) The gain field of the same unit. The gain field is the responsiveness of the unit to a target in the same location in the RF plotted against a 8 × 8 grid of eye positions, spaced by 10°. The gray level and the size of the squares in the 2-D gain field represent the activation of the unit. (C) The histogram of direction differences of the gain field and the receptive field of all hidden units in Model 1-3. The direction of a receptive field was calculated as the vector direction from the center of the input map to the center of the RF. The difference between the directions of the gain field and the RF of a unit was computed for every hidden unit. Most units have direction differences close to 180°, i.e., the ''opposite'' gain field structure.

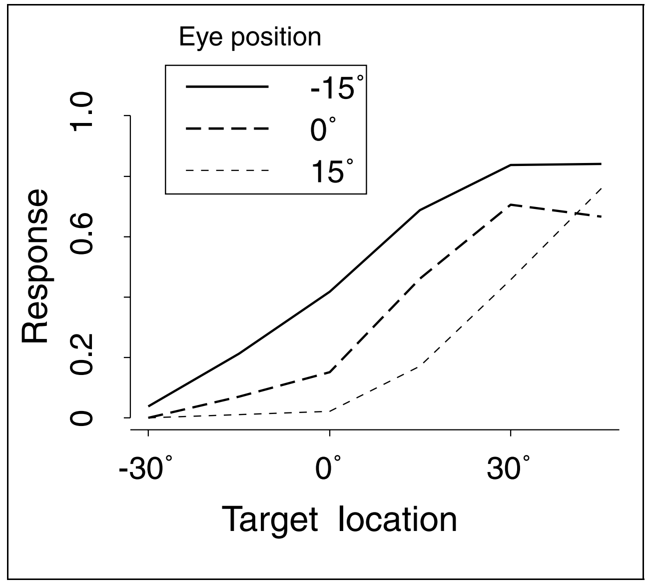


Figure 5. An example of RF shifts induced by eye-position modulation. The responses of a typical hidden unit in Model 1-3 are plotted against five auditory target locations. The vertical axis represents response magnitudes and the horizontal axis represents target locations. Tuning curves of a unit obtained at three eye positions are dashed differently. The tuning curves are distorted by eye-position modulations producing apparent horizontal shifts.

RESULTS

Level-1: No Visual-Auditory Integration

Model 1-1 and Model 1-2—No Coordinate Transformations

The hidden units exhibited localized RFs after the training was completed. The responses changed little with eye position; no hidden units developed GFs. Thus, the eye-position information was not combined with sensory information in the hidden layer. These results are the same for Model 1-1 and Model 1-2, since the two models have similar structures.

In these two models, since eye position had little effect on the hidden units, the VRFs should be anchored to the eyes while the ARFs should be anchored to the head-centered map. We examined the RF shift ratios for ARFs and VRFs. Figure 3A, B shows examples of 1-D RF tuning curves as a function of eye position. The RF curves are the unit's responses to targets presented at different angular locations of the input map. Three curves were obtained at the three eye positions (−20°, 0°), (0°, 0°), and (20°, 0°). Figure 3C, D shows histograms of RF shift ratios for Model 1-1 and Model 1-2, respectively. For each hidden unit, the RF shifts along the horizontal axis with horizontal eye position displacements and the RF shifts along the vertical axis with

Downloaded from http://direct.mit.edu/jocn/article-pdf/12/4/601/... on 18 May 2021

vertical eye position displacements were computed. The horizontal shift ratio and the vertical shift ratio for a unit are then plotted in the same histogram. Therefore, the histograms of Figure 3C and D each contain 40 shift ratios for the 20 hidden units. The same holds true for all the shift ratio histograms throughout this report. The mean VRF shift ratios for Model 1-1 is 0.97, whereas the mean of ARF shift ratios for Model 1-2 is 0.19. The deviation of shift ratios from zero or one is due to edge effects of the large RFs.

Model 1-3—Head-Centered to Eye-Centered Coordinate Transformation

This model transforms inputs of auditory target locations encoded in head-centered coordinates onto the eye-centered representation. After approximately 2,000 training trials, the network learned to map target locations onto the eye-centered output correctly. The hidden units developed localized ARFs and GFs for eye position. Figure 4 shows the RF and GF of an example unit. The RF size is about a quarter of the modeled input map (Figure 4A). The RF centers were all located at peripheral locations in the input map. Thus, none of the hidden units was nondirectional with the RF center located in the center of the input map. The GFs monotonically increased in a particular direction (Figure 4B).

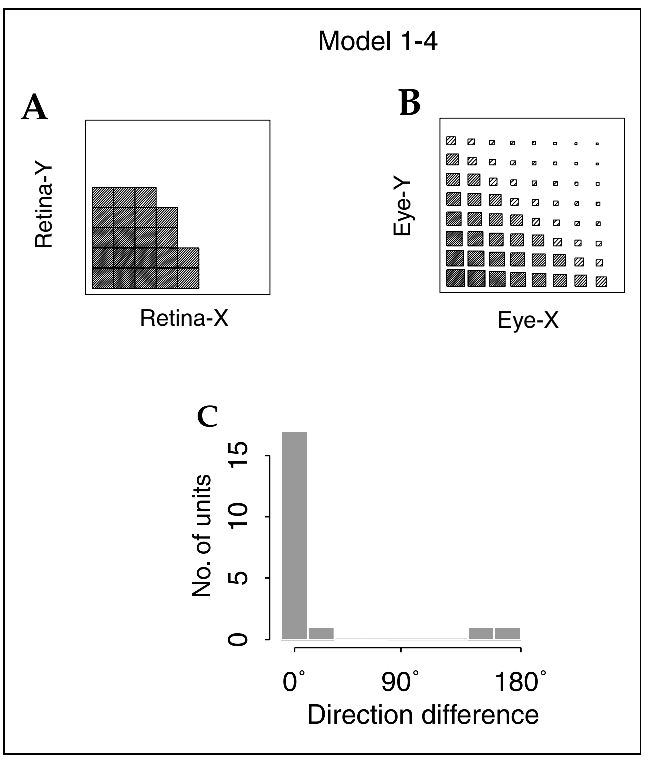


Figure 6. Receptive fields and gain fields of hidden units in Model 1-4. The layout of the figure is the same as that of Figure 3. Most hidden units have direction differences close to 0°, i.e., the ‘‘aligned’’ gain fields.

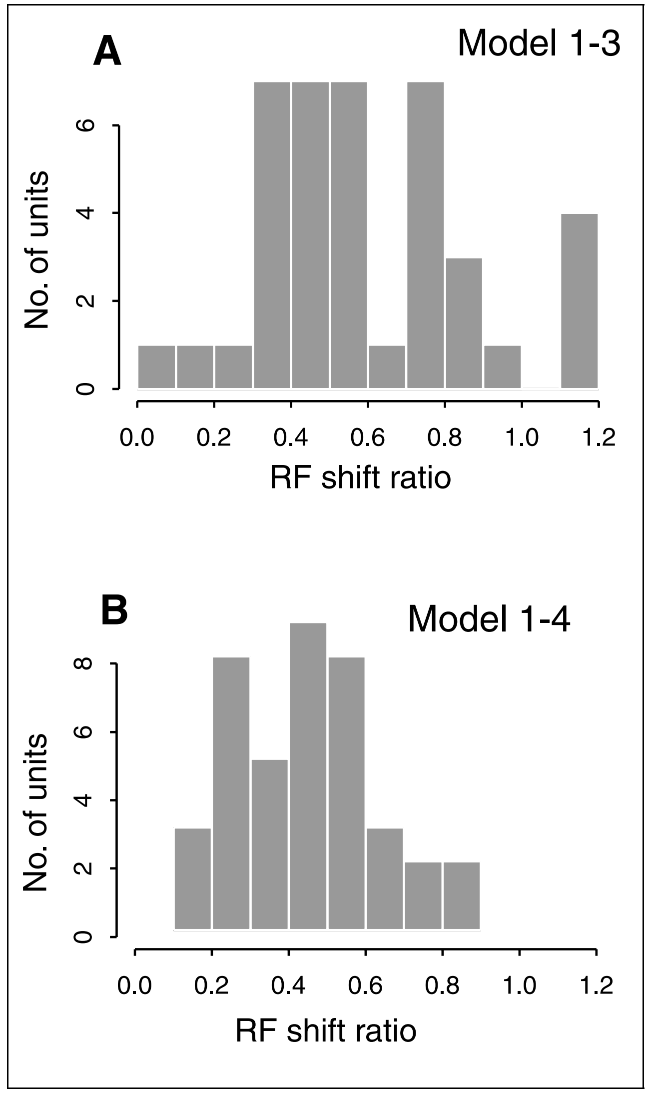


Figure 7. Histograms of coordinate indices of all hidden units in (A) Model 1-3 and (B) Model 1-4.

An interesting observation is that the unit had the GF approximately opposite to the direction of the unit's RF. This is the case for the majority of the hidden units. The histogram of the GF–RF direction differences for all hidden units is shown in Figure 4C. About 80% of the units have a direction difference close to 180°. Thus, the hidden layer of Model 1-3 essentially employs an ‘‘opposite GF’’ mechanism.

The ARF centers of the hidden units shifted with eye position. Figure 7A shows the histogram of ARF shift ratios for all the hidden units. The histogram included both horizontal and vertical shift ratios for each unit. The distribution of the ratios was broad with the mean ratio at 0.5, indicating that the ARF shifts was smaller than the shift of eye position. Further examinations showed that the responses of these units were selective, in varying degrees, for both eye-centered motor error and head-centered target location. Therefore, hidden units in Model 1-3 encode auditory inputs neither exclusively in

Downloaded from http://direct.mit.edu/jocn/article-pdf/12/4/601/1758996/089892900562372.pdf by guest on 18 May 2021

eye-centered frames nor exclusively in head-centered frames, but in intermediate coordinate frames.

The responses of nearly all the hidden units were modulated by eye position. This modulation resulted in distortions of ARF profiles, which lead to shifts of the RF centers. Two factors contribute to the shifts. (1) The sigmoidal integration of eye position and target location causes nonlinear modulation of the responses. The eye-position modulation is very effective when the response is weak, and the modulation saturates when the response is strong. These effects lead to the distortion of ARFs. (2) Eye position modulation introduces RF shifts. The RFs of the hidden units are usually very large so that the modulation is usually asymmetric on the RFs. This is because most RFs are located in the periphery and only one flank of the RF is in the region tested. Figure 5 illustrates an example of an ARF shift. The three curves in Figure 5 show the responses to target locations measured at three eye positions for a typical hidden unit, with the horizontal axis indicating target locations in head-centered coordinates and the vertical axis representing responsiveness. The response magnitudes are modulated by eye position. The activity saturates at the edge of the RF. The result of such nonlinear modulation is a shift of the RF. GF modulation, thus, provides a simple way to achieve a partial coordinate transformation without requiring a complicated shift circuit.

Model 1-4—Eye-Centered to Head-Centered Coordinate Transformation

This is the Zipser–Andersen model. The network learned to map visual inputs onto the head-centered output layer. Like those in Model 1-3, the hidden units in Model 1-4 developed localized VRFs and GFs. In contrast to Model 1-3, most hidden units of Model 1-4 had GFs tuned for the same direction as the unit's VRF. Figure 6A, B shows an example. The GF–RF direction differences for all hidden units are illustrated in the histogram of Figure 6C. Most units had direction differences close to zero. The VRFs of hidden units in Model 1-4 are not strictly anchored to the eyes. The VRF shift ratios shown in Figure 7B indicate that RF centers of most hidden units have a shift ratio smaller than 1.0. The RF centers of the hidden units were affected by the change of the eye position because of the gain modulation.

The hidden units in Model 1-1 and Model 1-2 did not exhibit GFs, because eye-position information is not used for these models in which no coordinate transformation occurs. The GF–RF relationship of the units in the Level-1 models can be classified into three categories: (i) opposite GF and RF directions, like those of model-3 units; (ii) aligned GF and RF directions, like those of Model 1-4 units; and (iii) GF and RF directions are neither significantly aligned nor significantly opposite. This type only constitutes a small portion of the hidden units in the Level-1 models. Aligned GF–RF were developed in order to transform signals from eye-centered frames to head-centered frames, while opposite GF–RF were to transform signals the other way around. These results indicate that the GF plays a crucial role in coordinate transformations.

Level 2: Integrated-Input Models

Model 2-1—Transformations of Visual and Auditory Inputs to Eye-Centered Coordinates

For each training cycle, a pair of eye position and visual target locations, or a pair of eye position and auditory target locations, were chosen at random as the inputs to the network. The expected output was the correspond-

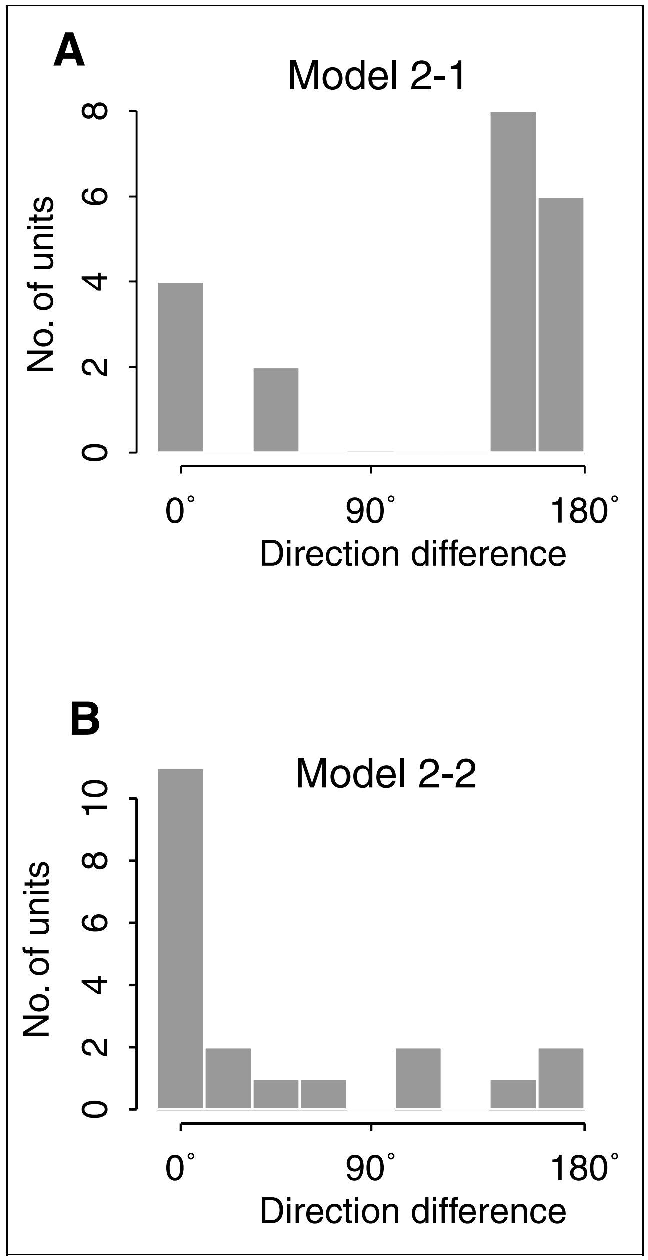


Figure 8. Histograms of direction differences of the hidden units in (A) Model 2-1 and (B) Model 2-2.

Downloaded from http://direct.mit.edu/jocn/article-pdf/12/4/601/1759339/089892900562363.pdf by guest on 18 May 2021

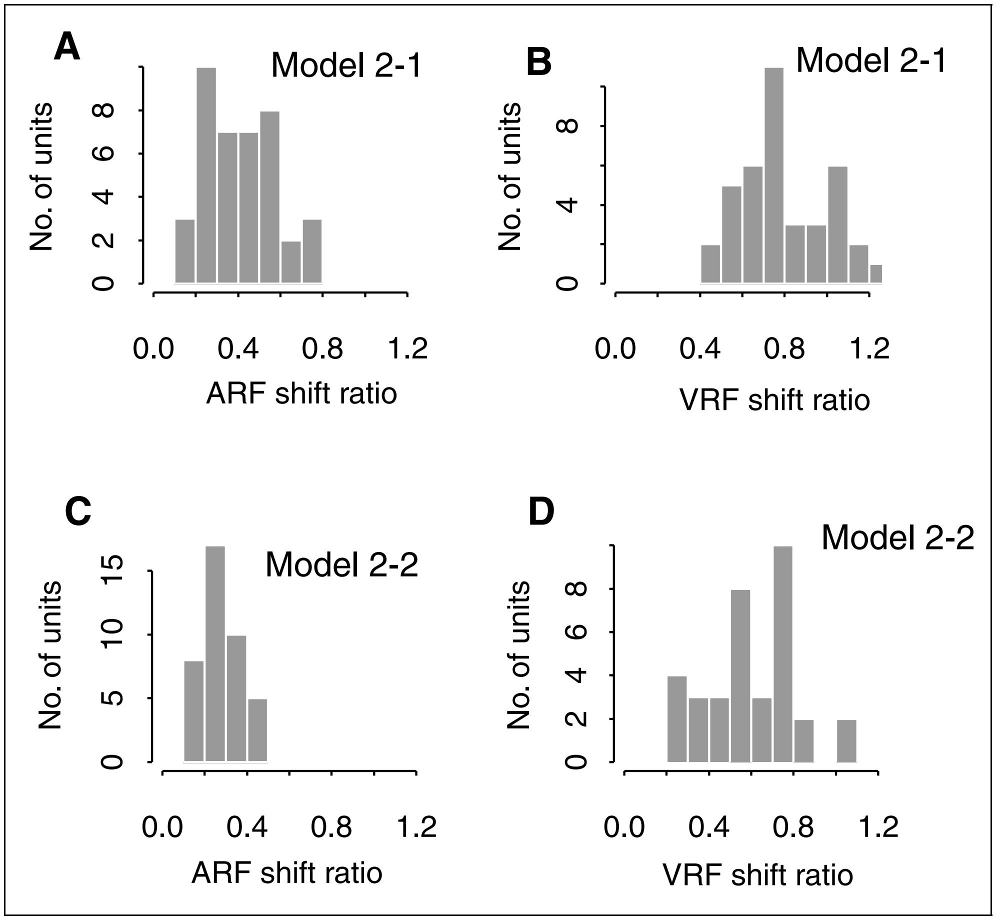


Figure 9. Histograms of RF shift ratios for Level-2 models. (A) ARF shifts of Model 2-1. (B) VRF shifts of Model 2-1. (C) ARF shifts of Model 2-2. (D) VRF shifts of Model 2-2.

ing eye motor error. The training process was stabilized after approximately 2,000 cycles. The model learned to map both visual and auditory inputs topographically onto the motor-error representation of the output layer.

The hidden units in Model 2-1 had similar properties to those in Model 1-3. (Both models perform the coordinate transformation from head-centered frames to eye-centered frames). After training, the hidden units developed localized ARFs and VRFs. The ARF and VRF of a unit usually aligned when the eye pointed to the central position. Most hidden units, like those in Model 1-3, had opposite GF–RF directions. Figure 8A shows the histogram of GF–RF direction differences for all the hidden units. Approximately 70% of the hidden units had direction differences close to 180°. These units were well tuned for motor errors and head-centered target locations. Among the six units whose GF–RF direction differences were close to zero, two of them had little response to any targets. The rest of the four units were selective to target location, but not selective to motor error. These results again demonstrate that the GF–RF direction difference is directly correlated to the coordinate transformation that the network performs.

The responses of the hidden units to auditory targets were found to be selective for both eye-centered and head-centered target location, suggesting "intermediate" coding of the hidden units. The ARF centers of the units shifted with eye position, but the shift did not completely compensate for the change of eye position, as shown in the histogram in Figure 9A. The mean of ARF shift ratios is about 0.5. Partial shifts of ARFs have been reported in PPC neurons (Stricanne et al., 1996).

Since visual targets were initially encoded in an eye-centered frame, no coordinate transformation was needed for visual signals. However, we found that responses of the hidden units to visual targets were not strictly anchored to the eye. Figure 9B shows the histogram of VRF shift ratios for all units. The mean of VRF shift ratios is 0.75. The shift of VRFs was due to eye-position modulation. Compared to the ARFs, the shift of VRFs was weaker. The peak of VRF profiles essentially stayed with eye position. This is because eye-position modulation was weaker for visual targets than for auditory targets. Detail examinations showed that slopes of the auditory GFs were steeper than the slopes of the visual GFs.

Model 2-2—Transformations of Visual and Auditory Inputs to Head-Centered Coordinates

This model is a mirror structure of Model 2-1. The training process was the same as that of Model 2-1 except that the expected output pattern was the head-

Downloaded from ... on 18 May 2021

centered representation. The hidden units in Model 2-2 have similar features to those in Model 1-4 because both models perform the coordinate transformation from eye-centered frames to head-centered frames. As we would expect from the results of Model 1-4, most hidden units in Model 2-2 have aligned GF–RF directions. Figure 8B shows the histogram of GF–RF direction differences for all the hidden units. Out of the 20 hidden units 15 (75%) had direction differences close to zero. The remaining units either had little response to any inputs or only exhibited tuning for motor error but not for head-centered target location.

The inputs of auditory targets were encoded in the head-centered frame; thus, no coordinate transformation was needed for auditory signals. However, the responses of the hidden units to auditory targets were still slightly shifted with eye position, like those VRFs in Model 2-1. Figure 9C shows the histograms of ARF shift ratios for all units. The mean of ARF shift ratios is 0.28. The mean of VRF shift ratios is 0.6, as shown in Figure 9D.

In summary, the Level-2 models are capable of bringing inputs from two different sensory modalities into the same coordinate frame, either eye-centered (Model 2-1) or head-centered (Model 2-2). The response properties of the hidden units in Model 2-1 and Model 2-2 are similar to those in Model 1-3 and Model 1-4. Most hidden units encode inputs in intermediate coordinate frames. The GFs play an important role in their intermediate coding. Stricanne et al. (1996) found that many PPC neurons indeed encoded auditory targets in intermediate coordinates and the auditory RFs of these neurons were partially shifted with eye position.

The Model of Level-3: Integrated Input–Output Model

At every training cycle, an eye position and either a visual or auditory target location were chosen randomly. There were two output patterns: the saccadic motor error in the eye-centered output map and the target location in the head-centered output map. The connection weights of the model were adjusted in order to reduce the performance errors of both outputs. It took about 5,000 training cycles for the training process to be

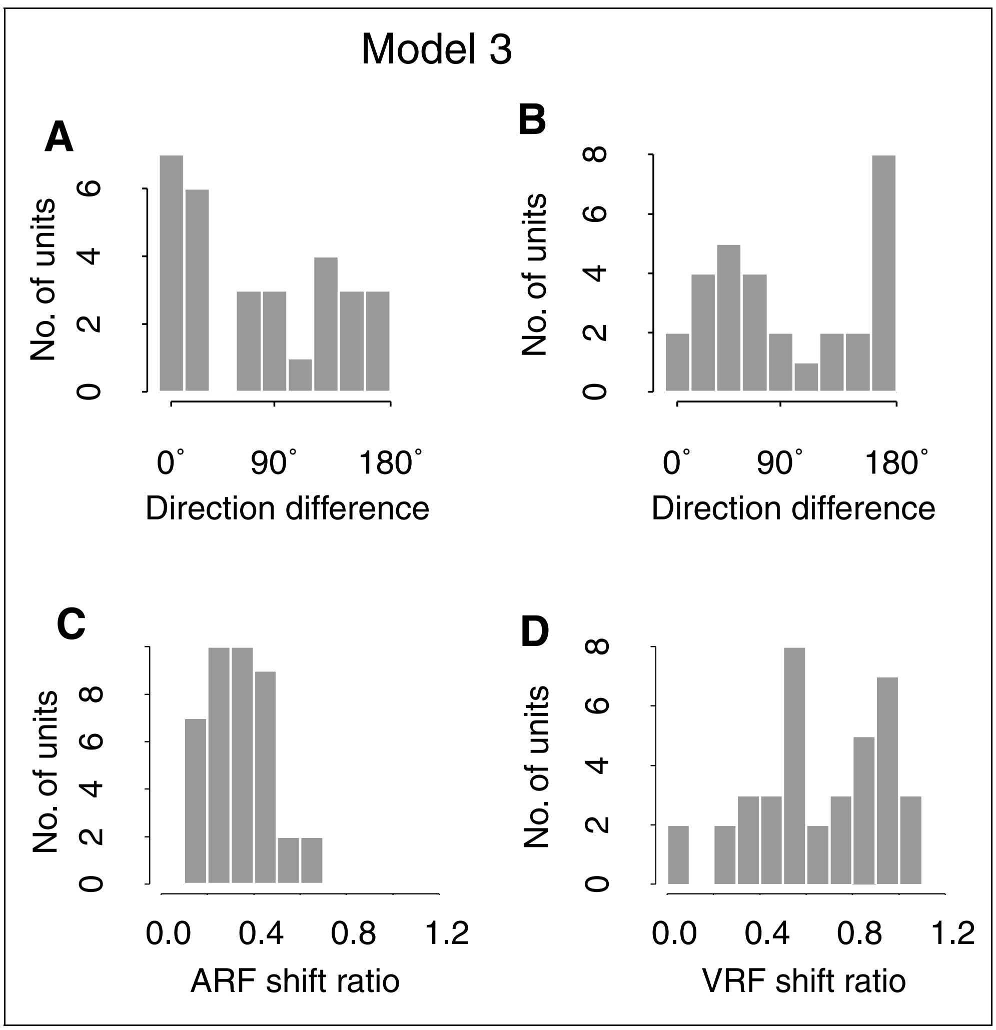


Figure 10. RF–GF relationship and RF shift ratios for Model 3. (A, B) Histograms of RF–GF direction differences for ARFs (A) and VRFs (B). (C, D) Histograms of ARF shift ratios (C) and VRF shift ratios (D).

Downloaded from http://direct.mit.edu/jocn/article-pdf/12/4/601/1758923/089892900562363.pdf by guest on 18 May 2021

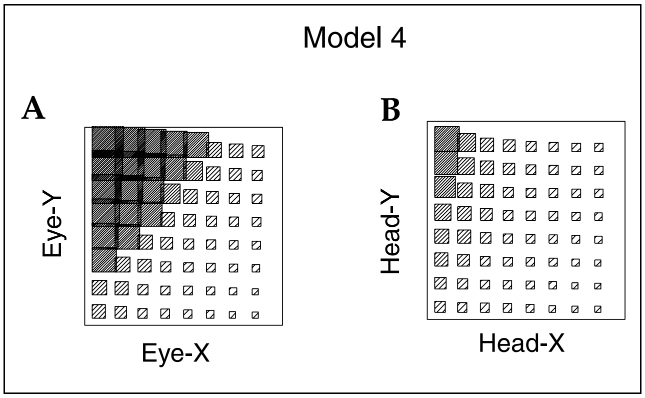


Figure 11. The gain fields for eye position and for head position of a typical hidden unit in Model 4. Both gain fields are aligned to the same direction; whereas the gain field for eye position (A) has a steeper slope than the gain field for head position (B).

stablized. The network learned to map visual and auditory inputs into both eye-centered and head-centered frames.

After training, the hidden units developed localized RFs and GFs. The RFs and GFs were similar to those in the Level-1 and Level-2 models. However, as shown in the histograms of Figure 10A and B, the RF–GF direction differences for all the hidden units spread from 0° to 180°. No explicit relationship between RF and GF could be derived. Since neurons with aligned RF and GF, opposite RF and GF, and neurons whose RF and GF direction differences are between 0° and 180° have all been observed in PPC (Stricanne et al., 1996), the PPC neurons were in better accordance with Model 3 units, which had a more diverse distribution of RF–GF direction differences than with Model 2 units. Therefore, the analysis of the population distribution of RF–GF relationship can provide information about what kinds of target representations exist in PPC.

The hidden units encoded both visual and auditory targets in intermediate reference frames. Figure 10C, D illustrates the histograms of ARF and VRF shift ratios. The ARF and VRF shifts of the hidden units in Model 3 are similar to those in Model 2-1 and Model 2-2. Thus, the measurement of RF shifts cannot differentiate neurons that provide a single-output representation from neurons that provide multiple-output representations.

The Model of Level-4: Integrated Input–Output Model with Head Location Information

In this model, head-position information is added into the input layer in the same manner as eye-position information. Besides the eye-centered and the head-centered output representations, the output layer of Model 4 has one more output map, which is the body-centered representation of targets: BE = H + E + V and BE = H + A. Thirty hidden units were used in Model 4. At each training cycle, an eye position, a head position, and a visual or an auditory target location were chosen at random as the inputs to the model. The connection weights of the model were adjusted according to the differences between activations of the three output maps and their expected outputs. After the training was stabilized, the network learned to produce correct output signals in eye-centered, head-centered and body-centered coordinate frames, respectively.

After training, the RF–GF relationship and coding properties of the hidden units in Model 4 were similar to those in Model 3. A new feature of Model 4 is that the hidden units developed spatial GFs for head position. Interestingly, the head-GF and eye-GF for the majority of the hidden units were in the same direction. The eye-GF and head-GF of a typical Model 4 unit are shown in Figure 11A and B. For a given unit, the slope of the eye-GF was, in general, slightly steeper than that of the head-GF. The average ratio of head-GF slopes over eye-GF slopes was 0.9. This agrees with the experimental results (Brotchie et al., 1995). For comparison, we trained the model to map inputs onto the body-centered output representation only. We found that the eye-GF slopes and the head-GF slopes were statistically the same. Thus, the steeper slopes of eye-GFs in Model 4 might reflect the fact that eye-position information was needed for all the three output representations, while head-position information was only needed for the body-centered representation.

DISCUSSION

Sensory signals of different modalities are initially encoded in different reference frames. On the other hand, the final actions of motor commands are also performed in different coordinate frames. For instance, saccades are made relative to the eyes, while the gaze shifts made with head movement are made relative to the head. Thus, the process of sensorimotor integration is involved in various types of coordinate transformations. How does the brain carry out these transformations? It is hard to imagine that the brain has many different circuits, each designed for one particular type of transformation. Some common principles for neurons involved in different types of transformations might exist. In this report, we examined several sets of coordinate-transformation models, which integrate inputs and outputs differently. Although these models differ in their structures and functions, we found that GF structure and intermediate coding are the typical features of the models for various types of coordinate transformations.

Structure of Models

Zipser and Andersen (1988) developed a model that combined visual and eye position signals, and mapped

Downloaded from http://direct.mit.edu/jocn/article-pdf/12/4/601/1759356/089892900562363.pdf by guest on 18 May 2021

them to a head-centered output layer. Krommenhoek et al. (1993) trained a backpropagation model to map input signals onto an eye-centered output layer. So far, there is no experimental evidence to exclude either one of these two output formats. Pouget and Sejnowski (1995) used the product of a Gaussian function of retinal location and a sigmoid function of eye position as a model of neurons. They demonstrated that a large set of such model parietal neurons could represent the position of an object in head-centered and eye-centered reference frames simultaneously. In this report, we assume that multiple representations of visual and auditory targets in different coordinates exist in the brain. The required coordinate transformations can be carried out by one or several groups of PPC neurons.

Four sets of models were analyzed in this report. We first examined the models without the integration of sensory inputs of different modalities. At this integration level, four separate sets of units carried out the transformations from visual or auditory inputs onto an eye-centered or head-centered representation. The hidden units in the four models of Level-1 received either visual or auditory target locations, but not both. These simple models allowed us to examine the nature of coordinate transformations. GF structure did not develop for the hidden units in Model 1-1 and Model 1-2, but did develop for the units in Model 1-3 and Model 1-4 in which coordinate transformations were required. Thus, GF structure is essential for the transformations either from eye-centered to head-centered coordinates or the other way around. As a result, the hidden units encoded inputs partially in eye-centered coordinates and partially in head-centered coordinates. In the Level-2 models, the hidden units received converging sensory inputs; two groups of the hidden units projected to an eye-centered or a head-centered output map, respectively. Simulations showed that such bimodal units were capable of bringing sensory information from different coding frames into the same-output coding frame. In Model 3, a single population of bimodal units combined visual, auditory, as well as eye position, information, and transformed target locations onto an eye-centered output map and a head-centered output map at the same time. The hidden units of this model developed gain-modulated visual and auditory RFs. The hidden units, as a population, have aligned, anti-aligned, and intermediate GF–RF direction differences. These properties are similar to those of PPC neurons (Stricanne et al., 1996). The similarities suggest that PPC is capable of proceeding multiple representations of space simultaneously. In Model 4, the hidden units had access to both eye position and head position. The model was trained to provide eye-centered, head-centered, and body-centered representations. After training, the hidden units successfully combined eye position and head position by having similar GFs for both of them. The experiments have found that more than half of PPC neurons exhibited similar GFs for eye position and head position (Brotchie et al., 1995), suggesting a possible body-centered representation of the space in PPC.

Coding Frame and RF Shift

PPC is involved in sensorimotor integration. This raises a question: Do PPC neurons encode targets in the coordinate frames of the sensory inputs or in the frames of motor outputs? For example, if a neuron encodes an auditory saccade target in the head-centered frame, its RF should be anchored to the head and should not change with eye rotation; if the neuron encodes motor errors of saccades, its auditory RF should be anchored to the eyes so that the RF shifts with the eyes. The RF shift ratio, defined as the induced shift of a neuron's RF divided by the corresponding shift in eye position, is used as a coordinate index throughout this report. Such a measurement has been used in experiments by Jay and Sparks (1987) and Stricanne et al. (1996).

In Model 1-1, where the hidden units simply relayed eye-centered visual inputs to eye-centered outputs, the coordinate indexes of the hidden units were close to 1.0. In Model 1-2, where the hidden units relayed head-centered auditory inputs to head-centered outputs, the indexes clustered near zero. In these models, the hidden units had no GFs. Therefore, the RFs did not shift with eye position. These results indicated that the measurement of the RF shift ratio was a reasonable estimate of a unit's coding frame. Figure 3 showed that the indexes of some units in Model 1-1 and Model 1-2 were not exactly equal to 0 or 1. The measurement errors were introduced when a RF was close to the edge of the input map.

Except in Model 1-1 and Model 1-2, all other models perform various types of coordinate transformations. The hidden units had GFs, which often caused partial shifts of RFs. The shift indexes of the hidden units did not cluster to 0 or 1. Instead, the indexes of all the hidden units in each model constituted various continuous distributions between 0 and 1. Thus, a series of intermediate frames were used by the hidden layer in order to perform the required coordinate transformations. The units, thus, coded targets neither exclusively in eye-centered coordinates nor exclusively in head-centered coordinates. These units employed the intermediate coding for targets, which is the result of the GF modulation of RFs by eye position signals. The advantage of such coding is that neurons contain information required for the output representations and also preserve information about the original sensory targets.

Throughout this report, intermediate coding was found in all the models that performed coordinate transformations, independent of specific transformation tasks. Therefore, this coding format may be a common feature of the sensorimotor systems. In fact, many neurons in PPC appeared to be intermediate coding neurons (Stricanne et al., 1996; Duhamel et al., 1997).

Downloaded from http://direct.mit.edu/jocn/article-pdf/12/4/601/1933596/089892900562363.pdf by guest on 18 May 2021

The similarities between the hidden units and a subset of PPC neurons suggests that PPC may be an intermediate stage in the process that brings different sensory modalities into the same coordinate frames by combining sensory signals with extraretinal signals.

Gain Fields

It has been well documented that most saccade-related neurons in PPC have retinotopic RFs that are gain-modulated by eye position. Zipser and Andersen (1988) showed that the hidden units combined eye position and visual target location by the means of GFs to form a head-centered output map. Results in our multimodel analysis allowed us to examine the intrinsic relationships between GF structure and coordinate transformations. In our simulations, only the hidden units of Model 1-1 and Model 1-2 in which no coordinate transformation occurred did not show GF modulation after training. Units in all other models developed GFs. Therefore, the coordinate transformations required by sensorimotor integrations are carried out through a GF structure.

The systematic relationship between the GF and the RF of a unit is very interesting. When the models were trained to map head-centered inputs onto eye-centered outputs, i.e., Output $= A - E$ (Model 1-3 and Model 2-1), the tuning directions of GFs were opposite to the RF directions. These units would maximally respond to the saccades that start from the favored GF area and end at the RF location; a rough selectivity for motor error is, thus, developed. On the other hand, the GFs and RFs of the hidden units were approximately aligned when the models mapped eye-centered inputs onto head-centered outputs, i.e., Output $= V + E$ (Model 1-4 and Model 2-1). These units roughly respond best to specific head-centered target locations irrespective of the initial eye position. Different types of coordinate transformations can, thus, be carried out by varying GF–RF relationships. In Model 3, the network has to transform inputs onto an eye-centered representation and a head-centered representation simultaneously, so that both opposite and aligned GF–RF are needed. As the result, the GF–RF direction differences of the hidden units fall into a continuous distribution between 0° and 180°. In Model 4, the hidden layer has access to signals of eye position and head position, and the output layer has three representations in different coordinate frames. The hidden units developed GFs for both eye position and head position. The head-GF and the eye-GF were often aligned, preserving a linear relationship of these two extraretinal signals. The aligned GFs for the eye position and the head position have been observed in PPC (Brotchie et al., 1995).

It is intriguing that different types of coordinate transformations can be carried simply through a GF mechanism. To achieve different coordinate transformations, all the network needs is to have different kinds of GF–RF relationships. This hypothesis is worth further investigation with electrophysiological experiments. Krommenhoek, Van Opstal and Van Gisbergen (1996) also showed that different combinations of the tuned directions to visual signals and to oculomotor signals yielded eye-centered or head-centered representations. Comparisons between experimental data and modelling results can highlight what kinds of coordinate transformations most likely occur in PPC. Experiments have found PPC neurons with opposite, aligned and intermediate GF–RF directions. These preliminary data are in better accordance with the results of Model 3 and Model 4 than with the results of the Level-1 and Level-2 models, suggesting the possibility of multiple representations in PPC.

CONCLUSION

We analyzed various possible models of sensorimotor integration in PPC. Our simulations showed that with GF structure and intermediate coding, a single network is capable of computing several representations of different coordinate frames. Although the available experimental results to date are not yet sufficient to be compared directly with the modeling results, the comparisons between the models suggest that RF–GF relationship could be used as an effective measurement in future experiments to examine the mechanisms of coordinate transformations in PPC.

Acknowledgments

This work was supported by NIH Grant No. EY05522. J. Xing was supported by the Del Webb Foundation fellowship. We thank B. Stricanne and J. Linden for many valuable discussions.

Reprint requests should be sent to Richard A. Andersen, Division of Biology, 216-76, California Institute of Technology, Pasadena, CA-91125, USA.

REFERENCES

Andersen, R. A., Essick, G. K., & Siegel, R. M. (1985). Encoding of spatial location by posterior parietal neurons. *Science, 25,* 456–458.

Andersen, R. A., Essick, G. K., & Siegel, R. M. (1987). Neurons of area 7a activated by both visual stimuli and oculomotor behavior. *Experimental Brain Research, 67,* 316–322.

Brotchie, P. R., Andersen, R. A., Snyder, L. H., & Goodman, S. J. (1995). Head position signals used by parietal neurons to encode locations of visual stimuli. *Nature, 375,* 232–235.

Duhamel, J. R., Bremmer, F., Ben Hamed, S., & Grof, W. (1997). Spatial invariance of visual receptive fields in parietal cortex neurons. *Nature, 389,* 845–848.

Helms Tillery, S. I., Flanders, M., & Soechting, J. F. (1991). A coordinate system for the synthesis of visual and kinesthetic information. *Journal of Neuroscience, 11,* 770–778

Jay, M. F., & Sparks, D. L. (1987). Sensorimotor integration in the primate superior colliculus: II. Coordinates of auditory signals. *Journal of Neurophysiology, 57,* 35–55.

Downloaded from ... on 18 May 2021

Krommenhoek, K. P., Van Opstal, A. J., Gielen, C. C. A. M., & Van Gisbergen, J. A. M. (1993). Remapping of neural activity in the motor colliculus: A neural network study. *Vision Research, 33,* 1287–1298.

Krommenhoek, K. P., Van Opstal, A. J., & Van Gisbergen, J. A. M. (1996). An analysis of craniocentric and oculocentric coding stages in a neural-network model of the saccadic system. *Journal of Neural Networks, 9,* 1497–1511.

Mazzoni, P., & Andersen, R. A. (1995). Gaze coding in the posterior parietal cortex. In M. Arbib (Ed.), *The handbook of brain theory and neural networks* (pp. 423–426). Cambridge: MIT Press.

Mazzoni, P., Bracewell, R. M., Barash, S., & Andersen, R. A. (1996). Spatially tuned auditory responses in area LIP of macaques performing delayed memory saccades to acoustic targets. *Journal of Neurophysiology, 75,* 1233–1241.

Mountcastle, V. B., Lynch, J. C., Georgopoulos, A., Sakata, H., & Acuna, C. (1975). Posterior parietal association cortex of the monkey: Command functions for operations within extrapersonal space. *Journal of Neurophysiology, 38,* 871–908.

Pouget, A., & Sejnowski, J. T. (1995). Spatial representations in the parietal cortex may use basis functions. *Advances in Neural Information Procedure, 7,* 157–164.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning internal representations by error propagation. In D. E. Rumelhart & J. L. McClelland (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition*, Vol. I (pp. 318–362), Cambridge: MIT Press.

Soechting, J. F., & Flanders, M. (1989a). Sensorimotor representations for pointing to targets in three-dimensional space. *Journal of Neurophysiology, 62,* 582–594.

Soechting, J. F., & Flanders, M. (1989b). Errors in pointing are due to approximations in sensorimotor transformations. *Journal of Neurophysiology, 62,* 595–604.

Soechting, J. F., Flanders, M., & Helms, Tillery, S. I. (1990). Transformation from head to shoulder-centered representation of target direction on arm movements. *Journal of Cognitive Neuroscience, 2,* 32–43.

Stricanne, B., Andersen, R. A., & Mazzoni, P. (1996). Auditory signals are transformed to eye-centered coordinates in posterior parietal cortex. *Journal of Neurophysiology, 76,* 2071–2075.

Synder, H. L., Brotchie, P., & Andersen, R. A. (1993). World-centered encoding of location in posterior parietal cortex of monkey. *Neuroscience Abstracts, 19,* 315.

Xing, J., Stricanne, B., & Andersen, R. A. (1994). A neural network model for sensorimotor transformation in macaque area LIP. *Neuroscience Abstracts, 20,* 143.

Zipser, D., & Andersen, R. A. (1988). A backpropagation programmed network that simulates response properties of a subset of posterior parietal neurons. *Nature, 331,* 679–684.

Downloaded from http://direct.mit.edu/jocn/article-pdf/12/4/601/1758990/089892900562363.pdf by guest on 18 May 2021