Two models for transforming auditory signals from head-centered to eye-centered coordinates

J. M. Groh and D. L. Sparks

Institute of Neurological Sciences and Department of Psychology, University of Pennsylvania, 3815 Walnut Street, Philadelphia, PA 19104, USA

Received March 12, 1992/accepted in revised form March 31, 1992

Abstract. Two models for transforming auditory signals from head-centered to eye-centered coordinates are presented. The vector subtraction model subtracts a rate-coded eye position signal from a topographically weighted auditory target position signal to produce a rate-code of target location with respect to the eye. The rate-code is converted into a place-code through a graded synaptic weighting scheme and inhibition. The dendrite model performs a mapping of head-centered auditory space onto the dendrites of eye-centered units. Individual dendrites serve as logical comparators of target location and eye position. Both models produce a topographic map of auditory space in eye-centered coordinates like that found in the primate superior colliculus. Either type can be converted into a model for transforming visual signals from retinal to head-centered coordinates.

1 Introduction

The deep layers of the primate superior colliculus (SC) contain a map of auditory space in oculocentric coordinates (Jay and Sparks 1984, 1987b). Units in this map respond best to sounds located at particular positions in space with respect to the direction of gaze. The receptive fields shift in position when the eyes move, maintaining an accurate representation of the location of auditory targets with respect to the eyes. Because they are encoded in a common coordinate framework, the representations of both visual and auditory targets remain in register and access a common motor circuitry for generating saccadic eye movements to these targets (Jay and Sparks 1987a).

Sounds are localized in space using interaural time and level differences as well as spectral cues. These parameters yield a head-centered representation of sound location which does not depend on eye position. Encoding the information in an eye-centered representation requires a transformation of coordinates. Two biologically plausible models for the implementation of this coordinate transformation are presented in this paper. Both models receive the same input and generate the same output, but they differ in network architecture, “cell” structure, the form of information coding in the intermediate stages, and the algorithms that they implement. A preliminary version of this work has appeared elsewhere (Groh and Sparks 1991).

2 The models

The transformation from head- to eye-centered coordinates is mathematically straightforward. Subtraction of a vector representing the position of the eyes with respect to the head from the location of the target with respect to the head yields the position of the target relative to the eyes. The first model, the vector subtraction model, implements this algorithm using neurons as computational elements. Subtraction occurs at inhibitory synapses. The second model, the dendrite model, executes the transformation through a convergence of eye position and craniocentric auditory target location information onto the dendrites of units in an oculocentric auditory map. The dendrites serve as comparators of eye position and target location information. The algorithm implemented by this model is analogous to a multi-dimensional look-up table cross-referencing craniocentric target locations and eye positions.

2.1 Vector subtraction model

Input to this model consists of a map encoding the location of an auditory target in craniocentric space and units encoding the horizontal and vertical components of eye position in the orbit (Fig. 1). At an intermediate stage, the eye position signal is subtracted from the craniocentric auditory target signal, yielding an oculocentric auditory target signal which is encoded in the firing rate of these intermediate units. This rate-code is converted into a place code for the output stage.
Fig. 1. The vector subtraction model. Excitatory input from the cranio-centric auditory map and inhibitory input from rate-coded eye position units converge on intermediate rate-coded oculo-centric units. The oculo-centric units excite both inhibitory interneurons and oculo-centric map units. The interneurons provide inhibitory input to the oculo-centric map. Notation: $x(a_l), y(a_e)$: horizontal and vertical position of the auditory target or receptive field with respect to the head, $x(a_s), y(a_s)$: position of the eyes with respect to the head, $x(a_l), y(a_e)$: position of the auditory target or receptive field with respect to the eyes using a combination of varying synaptic weights and inhibitory interneurons.

The cranio-centric auditory map consists of a 5 by 5 array of units. The units have non-overlapping receptive fields covering space within $\pm 20^\circ$ (horizontal and vertical) in the oculomotor range of the animal. The units fire if a target is present in the receptive field and are silent otherwise. The eye position units encode eye position in their rates of firing. Two of these units encode the horizontal component of eye position, one with a positive slope

$$f(x) = x + k$$

and the other with a negative slope

$$f(x) = -x + k$$

where $k$ is a constant ($+30^\circ$) which keeps the expression in the positive domain. The other two units encode the vertical component in a similar manner.\(^1\) All four eye position units are active for any position of the eyes in the orbit.

\(^1\) Because of the binary response functions of the non-rate coding units in the model, eye positions are constrained to be $10^\circ$ increments within $\pm 30^\circ$ of $0^\circ$.

Fig. 2. The connections of the model. All units and synapses are shown. Inhibitory synapses are depicted as T-shaped and excitatory synapses are fork-shaped. Notice that the lines depicting separate axonal projections from different units in the cranio-centric map are drawn as joining together with one another due to space considerations. For the example simulation, the target is located at $x(a_l) = 10^\circ$, $y(a_e) = 0^\circ$, where $x(a_l)$ and $y(a_e)$ are horizontal and vertical position of the auditory target with respect to the head. The eyes are directed at $x(a_s) = -20^\circ$, $y(a_e) = -10^\circ$, where $x$ and $y(a_e)$ are the position of the eyes with respect to the head. The darkened unit in the cranio-centric auditory map whose receptive field is located at $(x(a_s) = 10^\circ, y(a_s) = 0^\circ)$ fires at a rate of 1. The synaptic weights of the projections from the cranio-centric map are shown. The eye position units firing rates are as shown: \(\text{left HEP} = 50\), \(\text{right HEP} = 10\), \(\text{up VEP} = 20\), \(\text{down VEP} = 40\). The intermediate units firing rates are: \(\text{left} = 0\), \(\text{right} = 40\), \(\text{up} = 20\), \(\text{down} = 0\). The synaptic weights for one quadrant of the oculo-centric map are shown. The darkened horizontal and vertical inhibitory interneurons fire at a rate of 1, inhibiting the corresponding rows and columns of the oculo-centric map. The darkened oculo-centric unit at $(x(a_s) = 30^\circ, y(a_s) = 10^\circ)$ is not inhibited and receives enough excitation to exceed threshold, so it fires at a rate of 1. The parameters $x(a_s)$ and $y(a_e)$ are the position of the auditory target with respect to the eyes.
which has been weighted according to target location, and the unit fires at a rate proportional to the horizontal or vertical component of the position of the target with respect to the eyes. The complementary slopes of the eye position units ensure that two of the four intermediate units will receive net excitatory input for all combinations of craniocentric target location and eye position. The computation is complete at this intermediate stage. However, the oculocentric auditory target information found in the primate SC is in the form of a map, a place-code. The conversion of the rate-code at the model's intermediate stage into the biologically indicated place-code remains to be done. The intermediate units project directly to units in the output map as well as to inhibitory interneurons which in turn synapse on the output map units. The output map is an array 9 units by 9 units. The units in this map have receptive fields identical in size to those of the oculocentric auditory map, and also cover the oculomotor range. The map's dimensions are necessitated by the size of the input map. The output map must be larger than the input map so that large differences between target location and eye position can be represented accurately: a target at 20° left relative to the head requires a 40° leftward saccade if the eyes are directed 20° to the right with respect to the head. Units fire if threshold is exceeded and are silent otherwise. Associated with this map are horizontal and vertical inhibitory interneurons. Each unit receives input from one horizontal and one vertical intermediate neuron, and one horizontal and one vertical inhibitory interneuron. The synapses of the connections from the intermediate units are weighted according to the position of the output unit in the map. More peripheral units have lower synaptic weights than more central units and require a higher rate of firing in the intermediate units to become active. When the input is great enough to drive a peripheral unit, the more central units are shut down by the inhibitory interneurons. This ensures that only one unit in this map will be activated by the presence of a target.

The pattern of synaptic weights and thresholds is given in general form in Table 1. Other patterns could work equally well: the essence of the model is in the architecture rather than the specific weights. To provide an intuitive understanding of how the model works, its activity for a particular target location and eye position will be described. Figure 2 shows the firing rates and synaptic weights for this example. The target is located at the right, at \( x(a_t) = 10^\circ \), \( y(a_t) = 0^\circ \). The eyes are directed down and left, at \( x(e_t) = -20^\circ \), \( y(e_t) = -10^\circ \). The location of the target with respect to the eyes is \( x(e_t) = 30^\circ \), \( y(e_t) = 10^\circ \), so the desired output of the model is for the oculocentric auditory map unit with a receptive field at \( (30^\circ, 10^\circ) \) to fire.

The unit in the craniocentric auditory map whose receptive field is located at \( (x(a_t) = 10^\circ, y(a_t) = 0^\circ) \) fires at a rate of 1. The left horizontal eye position unit fires at a rate of 50, while the right horizontal eye position unit fires at a rate of 10. The up vertical eye position unit fires at a rate of 20, and the down vertical eye position unit fires at a rate of 40. The input to the intermediate units is the sum of the product of the synaptic weights and the firing rates of the presynaptic units. The intermediate units are linear threshold units whose firing rates are the weighted sum of the input minus the threshold. The thresholds of the intermediate units are \(-10\). The left horizontal intermediate unit receives input from the active craniocentric auditory unit with a weight of 20 and from the left horizontal eye position unit with a weight of \(-1\). The total input to this unit is

\[
(20)(1) + (-1)(50) = -30
\]

which is less than its threshold of \(-10\), so the unit does not fire. The active craniocentric auditory unit projects to the right horizontal intermediate unit with a weight of 40, and the right horizontal eye position unit projects with a weight of \(-1\). The total input is

\[
(40)(1) + (-1)(10) = 30
\]

The unit fires at a rate of 40 (input + threshold of \(-10\)). The auditory unit projects to both vertical intermediate units with a weight of 30. The up vertical intermediate unit receives a total input of

\[
(30)(1) + (-1)(20) = 10
\]

and fires at a rate of

\[
10 - (-10) = 20
\]

| Table 1 | Parameters of the vector subtraction model. The constant \( k = 30 \); \( x(a_t), y(a_t) \) = receptive field position with respect to the head; \( x(e_t), y(e_t) \) = eye position with respect to the head; \( x(a_t), y(a_t) \) = receptive field position with respect to the eyes. \( \Sigma w_a = \text{sum of } (\text{weights})(\text{firing rates}) \) |
|---------|-------------------------------------------------|----------------|----------------|
| Unit | Synaptic weights | Threshold | Firing rate function |
| Craniocentric auditory units | To intermediate units: Horizontal: \( k \pm x(a_t) \), Vertical: \( k \pm y(a_t) \) | NA | \( 1 \) if target present, 0 otherwise |
| Eye position units | To intermediate units: \(-1\) | NA | Horizontal: \( k \pm x(e_t) \), Vertical: \( k \pm y(e_t) \) |
| Intermediate units | To inhibitory interneurons: \( l/(\pm x(a_t) + 20) \), To oculocentric units: \( l/(\pm x(a_t) + 10) \) | \(-10\) | \( \Sigma w_a \) - threshold |
| Inhibitory interneurons | To oculocentric units: \(-11\) | 1 | \( 1 \) if \( \Sigma w_a \geq \text{threshold} \), 0 otherwise |
| Oculocentric auditory units | | 2 | \( 1 \) if \( \Sigma w_a \geq \text{threshold} \), 0 otherwise |
The down vertical intermediate unit received a total input of
\[(30)(1) + (-1)(40) = -10,\]  
and is silent since
\[-10 - (-10) = 0.\]

Units in the upper right quadrant of the oculocentric auditory map receive input from both the right horizontal intermediate unit and the up vertical intermediate unit, which are the two active intermediate units. Only the units with synaptic weights such that the weighted sum of the input exceeds threshold will fire. All units in this map have a threshold of 2. The minimum combination of weights that meet this requirement are a weight of 1/40 for the horizontal input and a weight of 1/20 for the vertical input. The unit with this set of weights is located at \((x(a_n) = 30^\circ, y(a_n) = 10^\circ)\) in the oculocentric auditory map. The horizontal inhibitory interneuron which synapses on that unit receives input from the same horizontal intermediate unit, but the weight is less than 1/40 and it does not fire. The same is true of the vertical inhibitory interneuron. The interneurons have thresholds of 1 since they receive input from only one intermediate unit.

Units located more centrally in the oculocentric auditory map have higher synaptic weights than the unit at \((30^\circ, 10^\circ)\), and thus are also receiving enough excitation to exceed threshold. However, these units are inhibited by the activity of the associated inhibitory interneurons. The conversion from a rate-code to a place-code is performed by letting each unit require a different amount of excitation to become active, and by shutting down the units that receive too much excitation with the inhibitory interneurons. A unit will only fire if the activity of the intermediate units is within a certain range: greater than its own threshold, but less than the thresholds of both the horizontal and vertical inhibitory interneurons associated with it.

The result of this pattern of weights and connectivity is that the only unit firing is the unit at \((30^\circ, 10^\circ)\). This is the position of the target relative to the eyes in this example.

Both positive and negative slope units encoding eye position and auditory target location with respect to the eyes at the intermediate stage are required. The operation performing the coordinate transformation is a subtraction, and neurons can only have positive firing rates. The complementary units ensure that a positive signal at the intermediate stage is always present if a target is present.

The model was simulated using NEXUS (Sajda and Finkel, 1992) on Sun Sparcstations. The inhibitory interneurons and craniocentric and oculocentric auditory map units were simulated using binary units, and the eye position units and intermediate units were simulated using linear threshold units.

2.2 Dendrite model

Like the vector subtraction model, the dendrite model receives input from a map of auditory space in craniocentric coordinates and from units encoding horizontal and vertical eye position in their rate of firing. However, the dendrite model eliminates the intermediate stage of a rate-coded signal of the oculocentric position of the auditory target. Instead, each unit in the craniocentric auditory map projects directly onto a separate dendrite of every unit in the oculocentric auditory map (momentarily setting aside edge effects). The local circuitry at each dendrite serves to compare points in head-centered space with current eye position and cause the unit to fire if a target is present at a particular position with respect to the eyes.

As in the vector subtraction model, the input craniocentric auditory map consists of a 5 by 5 array of units, whose non-overlapping receptive fields span \(\pm 20^\circ\) of the oculomotor range of the animal (Fig. 3). The eye position inputs are also identical to those used in the vector subtraction model. A 9 by 9 array of units forms the oculocentric auditory map. Each unit in this map has one dendrite per presynaptic craniocentric auditory unit. These units receive input from all craniocentric units that could potentially signal a target in their receptive fields. For example, a craniocentric unit whose receptive field lies at the extreme right side of the oculomotor range does not project to oculocentric units whose receptive fields lie to the left of the direction of gaze. The dendrites of the oculocentric units also receive input from the horizontal and vertical eye position units.

**Fig. 3.** The dendrite model. Craniocentric auditory map units and rate-coded eye position units project to inhibitory interneurons and the dendrites of oculocentric map units. The inhibitory interneurons also project to the oculocentric units, with each oculocentric unit receiving input from many interneurons. The synapses made by craniocentric units are excitatory, the synapses of the eye position units and interneurons are inhibitory.
and from inhibitory interneurons. The inhibitory interneurons receive input from single craniocentric units and either horizontal or vertical eye position units.

Figure 4 shows the dendritic circuitry of a sample oculocentric unit. A typical dendrite receives excitatory synapses from one craniocentric unit at the tip. Inhibitory synapses from a horizontal and a vertical eye position unit are located more proximally on the dendrite. The horizontal inhibitory interneuron associated with the dendrite receives excitation from the same craniocentric unit and inhibition from the horizontal eye position unit. The vertical inhibitory interneuron receives excitation from that craniocentric unit and inhibition from the vertical eye position unit. They synapse onto the proximal dendrite and produce shunting inhibition if active. If a target is present which activates the craniocentric auditory unit, the dendrite and interneurons will both receive excitation. The weights of the inhibitory synapses are set so that if the eye position signals are above a certain limit, the inhibition from the direct projection onto the dendrite from the eye position units will outweigh the excitation, and if they are below a certain limit, the inhibitory interneurons will become active and shut down the dendrite. Between these limits, the excitation will outweigh the inhibition and drive the oculocentric unit. The region between these limits corresponds to the receptive field.

Table 2 shows the thresholds, activation functions, and the rules which describe the pattern of synaptic weights in the model. Again, many other patterns could work equally well. The general solution to the model for this pattern is shown in the appendix.

To demonstrate how the model functions let us examine what happens at one dendrite of an oculocentric unit for a single auditory target as the eye moves through three different positions. The oculocentric unit has a receptive field up and to the right, at (10°, 10°). The dendrite of interest receives input from the craniocentric unit with a receptive field at (0°, 0°). Figure 5 shows the weights for all the synapses. Let the auditory target be straight ahead at (0°, 0°) with respect to the head, in the receptive field of the craniocentric unit. In the first case, let the eyes be directed straight ahead at (0°, 0°) as well. The horizontal and vertical eye position units both fire at a rate of 30. The input to the horizontal inhibitory interneuron is

\[(1)(1 + (-1/10))(30) = -2, \tag{9}\]

<table>
<thead>
<tr>
<th>Unit</th>
<th>Projection from</th>
<th>Synaptic weights</th>
<th>Threshold</th>
<th>Firing rate function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oculocentric dendrite</td>
<td>Craniocentric auditory unit</td>
<td>2</td>
<td>0</td>
<td>1 if (\Sigma w_a \geq) threshold, 0 otherwise</td>
</tr>
<tr>
<td>if (x(a_h) \geq 0)</td>
<td>HEP with firing rate (x(a_h) + k)</td>
<td>(-1/(k + x(a_h) - x(a_v)))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>if (x(a_h) &lt; 0)</td>
<td>HEP with firing rate (-x(a_h) + k)</td>
<td>(-1/(k - x(a_h) + x(a_v)))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>if (y(a_v) &lt; 0)</td>
<td>VEP with firing rate (y(a_v) + k)</td>
<td>(-1/(k - y(a_v) - y(a_h)))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vertical inhibitory interneuron</td>
<td>VEP with firing rate (-y(a_v) + k)</td>
<td>(-2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Horizontal inhibitory interneuron</td>
<td>Horizontal and vertical inhibitory interneurons</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>if (y(a_v) \geq 0)</td>
<td>Craniocentric auditory unit</td>
<td>1</td>
<td>0</td>
<td>1 if (\Sigma w_a \geq) threshold, 0 otherwise</td>
</tr>
<tr>
<td>if (y(a_v) &lt; 0)</td>
<td>HEP with firing rate (x(a_h) + k)</td>
<td>(-1/(k + x(a_h) - x(a_v) - 10))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vertical inhibitory interneuron</td>
<td>HEP with firing rate (-x(a_h) + k)</td>
<td>(-1/(k - x(a_h) + x(a_v) - 10))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>if (y(a_v) \geq 0)</td>
<td>VEP with firing rate (y(a_v) + k)</td>
<td>(-1/(k + y(a_v) - y(a_h) - 10))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>if (y(a_v) &lt; 0)</td>
<td>VEP with firing rate (-y(a_v) + k)</td>
<td>(-1/(k - y(a_v) + y(a_h) - 10))</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
so it does not fire. The vertical inhibitory interneuron also receives input totalling $-2$ and is silent. The input to the dendrite is

$$\text{(2)}(1) + (-1/20)(30) + (-1/20)(30) = -1 .$$

Threshold is not reached and the unit does not fire. This is as it should be, since the target is not in the receptive field of the unit.

If the eyes move down and to the left to $(-10^\circ, -10^\circ)$, the target enters the receptive field of the unit. The eye position units both fire at a rate of 20. In this case, the input to the horizontal and vertical inhibitory interneurons is

$$\text{(1)}(1) + (-1/10)(20) = -1 ,$$

and they do not fire. The input to the dendrite is

$$\text{(2)}(1) + (-1/20)(20) + (-1/20)(20) = 0 .$$

The threshold is 0, so the unit fires.

When the eyes move still further down and to the left, to $(-20^\circ, -20^\circ)$, the target is once again out of the receptive field. Both of the eye position units fire at a rate of 10. Now the input to the horizontal and vertical inhibitory neurons is

$$\text{(1)}(1) + (-1/10)(10) = 0 ,$$

and they both fire. The input to the dendrite is

$$\text{(2)}(1) + (-1/20)(10) + (-1/20)(10) + (-2)(1) + (-2)(1) = -3$$

which is less than 0, and the unit does not fire.

3 Discussion

3.1 Vector subtraction model

A major asset of the vector subtraction model is that it is extremely parsimonious. It uses the minimum neural circuitry necessary to accomplish the task. The computation is performed at the intermediate synapses, with subsequent processing serving only to convert the information to the format that has been found in the brain. This model would be simpler still if cranio-centric target position were already encoded in the form of neural firing rates. Units encoding the horizontal or vertical component of cranio-centric target location would form excitatory synapses on the intermediate units and the eye position signal would be subtracted using inhibition. The graded weighting scheme required for input in map form would not be necessary. Candidates for the auditory input will be discussed in a subsequent section.

Other mechanisms could serve the same function as the inhibitory interneurons in the conversion from the intermediate rate-code to the final place-code. Biophysical membrane properties could make units fire only when the input is within a certain "resonant" range. Chloride or potassium channels activated by higher concentrations of neurotransmitters or by large depolarizations could serve the purpose. Such a biophysical mechanism may be more likely since it would require a smaller investment in neurons by the brain to accomplish what may be a common task.

A feature inherent to the vector subtraction model is that it can only compute the oculo-centric position of one auditory target at a time. Only the four intermediate units transmit information from the input to the output, and these are all required to provide a unique specification of the position of a single target. If two targets were represented at the input cranio-centric auditory map, the intermediate stage would sum the resulting activity, and the output would be too large. A prediction can be made from this restriction: if the model is correct, this neural circuitry must lie subsequent to target selection circuitry in the brain. The auditory target must be chosen as a potential target of a saccade before it is transformed into oculo-centric coordinates. Recording from the oculo-centric map during simultaneous presentation of two targets should show only one site of activity on the map. Two-point stimulation in the input map would result in erroneously large signals at the intermediate stage and activation of more peripheral units in the output map. Deactivation of small regions of the input map using
Lidocaine would cause erroneously small signals and activation of more central units in the output map. A small modification of the model could make these errors more biologically plausible. If the intermediate units could extract the weighted average of the activity of the input map instead of the weighted sum, the output would be the vector average of the two targets. Vector averaging occurs in other oculomotor contexts such as two-point collicular stimulation, lidocaine deactivation of the SC, and various behavioral paradigms (Robinson 1972; Robinson and Fuchs 1969; Schiller et al. 1979; Lee et al. 1988; Becker and Jurgens 1979; Van Opstal and Van Gisbergen 1990).

3.2 Dendrite model

The dendrite model was designed to solve the single target problem. Since there is no limiting intermediate stage, multiple targets represented on the input map are preserved on the output map. The individual dendrites serve as independent logical devices (Shepherd and Koch 1990). The function of each dendrite can be described as a conditional: if excitation is received from the cranioconicentric unit and the eyes are within a certain range then pass the excitation to the soma. The soma performs a logical or across the dendrites and the unit fires if any are providing excitation. Even when there are multiple auditory targets, no more than one dendrite can excite the soma of a given unit because the eyes can only be at one position at a time. The other targets will excite different units. If no targets are present, the activity of the eye position units will inhibit all the oculocentric units and inhibitory interneurons.

This model could be implemented using simpler connectionist-style units, but many more of them would be required. Each dendrite would have to be modelled using one connectionist unit. Using the same scale, the input map of dimension n = 5 would contain n^2 = 25 cranioconicentric units, the output map would contain (2n - 1)^2 = 81 units, and n^4 = 625 units would replace the dendrites. The number of horizontal and vertical inhibitory interneurons is the same: n^3(n - 1)^2 or 400 each. (There are fewer inhibitory interneurons than dendrites because dendrites receiving input from the extremes of the potential cranioconicentric locations within the oculocentric receptive field do not need the upper cut-off that the inhibitory interneurons provide.)

Using dendrites thus greatly reduces the number of neurons necessary to solve the problem with this algorithm. Real neurons in the brain have extensive dendritic arbors, and models which exploit this complexity may approach more closely the solutions used by the brain for particular problems.

It may be unrealistic to expect units in the cranioconicentric map to project to all the units in the oculocentric map. In fact, this convergence could occur in several steps. Intermediate units would receive input from subsets of cranioconicentric map units. They would be selective for combinations of target position and eye position, but they could only respond to targets in a subregion of cranioconicentric space. Units with the same oculocentric selectivity for different regions of cranioconicentric space would converge on the same output unit, providing that unit with an oculocentric receptive field sensitive to targets anywhere in cranioconicentric space. If one recorded from the intermediate units, they would have oculocentric receptive fields gated by cranioconicentric target location, which could appear to be cranioconicentric receptive fields gated by eye position.

3.3 Biological basis

The technique employed in designing these models was to determine an algorithm that could realistically be implemented using neural elements, and then to develop an architecture and pattern of synaptic weights to carry out the algorithm. Modeling in other systems often involves training networks using back-propagation to solve problems for which the algorithm is not known. However, for simpler problems this technique is not necessary and can even obscure the issues by resulting in a model whose algorithm is not readily comprehensible. Incomprehensible models of the brain do not particularly aid in understanding its function, though they may, of course, be correct.

Nonetheless, an associative or Hebbian mechanism could certainly be used to establish the synaptic weights of either the vector subtraction or the dendrite model. An instructor signal is readily available: oculocentric visual input to the output layer could "teach" the auditory input. The majority of auditory cells in the primate SC also respond to visual input. In barn owls the role of vision in calibrating sound localization has been well documented by Knudsen and colleagues (Knudsen and Knudsen 1985; Knudsen and Brainard 1991; Knudsen et al. 1991).

These models are anchored to the known physiology at the level of the output network, which is intended as the auditory map in the deep layers of the primate SC. The finer details of this map were not included in the model. For example, the auditory receptive fields tend to be broader vertically than horizontally, the overall topography of the motor map is logarithmic in its representation of saccade amplitude, and the visual input was omitted (Jay and Sparks 1987b; Robinson 1972). These features were sacrificed for the sake of simplicity.

The sources of input were chosen for their biological plausibility. Neurons carrying rate-coded eye position signals similar to those used in the models have been found in a variety of places in the brain, including the brain stem gaze centers (Luschei and Fuchs 1972; Keller 1974), frontal eye fields (Bizu 1968), nucleus prepositus hypoglossi (McCrea and Baker 1980), the flocculus of the cerebellum (Noda and Suzuki 1979), and posterior parietal cortex (Andersen et al. 1987). The firing rate of these neurons is typically linearly related to eye position. This information is thought to be derived from either a copy of the motor command to move the eyes or proprioceptive feedback from the extraocular muscle spindles. No evidence of a place-code of eye position, with neurons having response
fields for eye position, has ever been reported. Since the
eyes can only be in one position at a time, a small
number of rate-coded units are sufficient to represent
eye position. The SC receives projections from areas
containing cells with a rate-code of eye position such as
parietal cortex, frontal eye fields, and cerebellum, and
projects directly to brainstem gaze centers and indirectly
to cerebellum (Sparks and Hartwich-Young

Support for the use of a neural map of auditory
space in head-centered coordinates is weaker. Few stud-
ies of auditory physiology have been conducted in the
primate. However, the more substantial body of data
on the cat can help identify plausible candidates for
such a map. This cross-species comparison can be
justified by the similarities between cat and primate SC.
Recent experiments suggest that the SC auditory map
in the cat, like the primate, is encoded in oculocentric
coordinates. Peck et al. (1992) and Vimal et al. (1989)
report auditory responses altered by eye position so as
to remain in register with the visual map.

The two most likely candidates for head-centered
auditory input to the SC are the inferior colliculus (IC)
and auditory cortex (Sparks and Hartwich-Young
1989). Data on a few units in the primate IC were
collected by Jay and Sparks (1987b) during the course
of their study on the SC. None of the five units they
studied were affected by eye position though they did
show some spatial selectivity with respect to the head.
In the cat, the auditory responses of the IC have main-
ly been studied in the anesthetized preparation. Semple
et al. (1983) and Aitkin and Martin (1990) have found
spatial selectivity, but the receptive fields all tend to be
centered on the acoustical axis of the contralateral
pinna rather than being distributed evenly across the
contralateral field. Thus, the IC does not appear to
contain a head-centered auditory signal in the form
called for by the models. The response latencies of SC
neurons to auditory stimuli (mean 44.8 ms) as com-
pared with IC neurons to the same stimuli (mean 28
ms) in the primate also suggests that this may not be
the primary pathway for this information (Jay and
Sparks 1987b).

By analogy with other sensory modalities, cortical
auditory input is a more promising candidate. Cooling
of a broad region of visual cortex reduced or elimi-
nated the visual responses of SC neurons in the cat
(Stein 1978). Cooling of SIV somatosensory cortex in
the anterior ectosylvian sulcus reduced the responses of
SC neurons to tactile stimuli in the cat (Clemo and
Stein 1986). Small lesions in the lateral suprasylvian
cortex produced visual neglect and reduced visual ac-
uity in the SC of cats (Hardy and Stein 1988).

A prominent source of cortical auditory input to the
SC is the anterior ectosylvian sulcus (AES). This
area contains visual, auditory and somatosensory cells
either in close proximity to or intermingled with each
other (Clarey and Irvine 1986, 1990; Clemo and Stein
1983; Olson and Graybiel 1987). Anterograde tracing
from this area shows that the deep layers of the SC are
a primary target (Segal and Beckstead 1989). In a
comparison with the other cortical auditory areas
(AI, AII, Fields A, P and VP), Meredith and Clemo
(1989) found that Field AES was the only one to show
a major anatomical projection to the cat SC. The
auditory function of this projection was confirmed by
sensory and electrical stimulation. Of collicular cells
identified as auditory, 73% were driven by electrical
stimulation in Field AES. Temporary deactivation of
AES by cooling caused a sample of these cells to cease
responding to auditory stimuli. Stimulation of other
cortical auditory areas did not drive SC neurons even at
currents greater than 600 μAmps. Given the anatomical
and electrophysiological evidence that these other areas
do not provide direct input to the cat SC, they will be
given no further consideration here.

The spatial selectivity of Field AES neurons to
free-field auditory stimuli has not been examined.
Clarey and Irvine (1986, 1990) studied the responses to
dichotic stimulation in anesthetized cats. Most neurons
were excited by monaural stimulation of either ear, and
in most of these the response to binaural stimulation
was less than the sum of the monaural responses. A few
cells could only be driven by binaural stimulation.
Noise bursts were more effective stimuli than tones.
Although these data do not provide direct evidence of
spatial selectivity in AES neurons, neither do they
rule it out.

If Field AES neurons are found to be spatially
selective in response to free-field stimuli, the coordinate
frame of these responses becomes important. It is pos-
sible that they have spatial receptive fields in eye-centered
coordinates rather than head-centered coordinates. If
so, the search for the site of the head-centered to
eye-centered coordinate transformation would shift to
the afferents of Field AES. If Field AES is found to
contain head-centered auditory information in some
form other than an explicit place-code, the input layer
of the models could be adapted to reflect this. Field
AES might also be found to contain an intermediate
stage of processing between head- and eye-centered
frames of reference, which could also be incorporated
into the models.

Both the dendrite and the vector subtraction models
generate predictions regarding the topography of the
connections between head-centered and eye-centered
auditory maps. Since an essential feature of the com-
putational problem is that all points in head-centered
space map onto all points in eye-centered space, signifi-
cant convergence and divergence is required. In the
vector subtraction model, the projections of the
cranio-centric map converge onto a small number of
intermediate units, whose projections then diverge
onto the large oculocentric map. In the dendrite
model, the convergence and divergence of projections
occur simultaneously.

The topography of the projection from AES to SC
in the cat was extensively examined by Meredith and
Clemo (1989) using both tracers and antidromic stimu-
lution. Injection of horseradish peroxidase into the SC
resulted in heavy labelling of Field AES, but little or no
labelling of the other auditory cortices. When small
amounts of different dyes were injected at two sites in the SC, the patterns of labelling in Field AES were intermingled. When small amounts of tritiated leucine were injected in Field AES, label was found throughout the intermediate and deep layers of SC. This corticocortical convergence and divergence was confirmed with stimulation. Single SC cells could be driven by orthodromic stimulation at many sites in Field AES but not from other auditory areas, and antidromic stimulation at many sites in the SC could activate the same site in Field AES. The extensive convergence and divergence of the projection is most consistent with the dendritic model, assuming that Field AES encodes head-centered auditory target location.

The relation of auditory association areas to the SC in the primate is less well studied, but Fries (1984) confirmed that auditory regions in the bank of the sylvian fissure and on the superior temporal gyrus project to the intermediate and deep layers of the SC. Leinonen et al. (1980) report that neurons in temporoparietal association cortex (Tpt) are broadly tuned for the spatial location of auditory stimuli. Neighbouring cells tend to prefer the same locations. In a few cases where head position was varied, the responses depended on the location of the sound with respect to the head rather than the body, but the effects of eye position were not examined. Cells responding to somatosensory and visual stimuli were also found in this region, lending support to the idea that it may be homologous to area AES of the cat.

3.4 Manipulation of the models

Further insight into the connection between the models and biology can be gained by manipulating or modifying various aspects of the models in ways that may occur experimentally. Some of these manipulations suggest possible interpretations for the observation that the auditory map in the SC retains broad spatial selectivity with respect to the head in the anesthetized cat (Middlebrooks and Knudsen 1984; Meredith and Stein 1986; Wickelgren 1971).

Changes in the eye position signal can have profound effects on the behavior of the models. If the eye position signal were completely removed by some selective lesion, the oculocentric auditory map of the dendrite model would become unresponsive to auditory stimuli because the inhibitory interneurons would silence any dendrite receiving excitation from the cranioocentric map. For the vector subtraction model, the site of activity in the oculocentric map in response to an auditory target would be shifted towards the periphery because the firing rates of the intermediate units would be too high. In addition, all the intermediate units would be active so units in all four quadrants of the oculocentric map would respond to the same auditory target. The oculocentric units would retain their spatial selectivity, but in head-centered not eye-centered coordinates. This pattern of activity would be dysfunctional for guiding saccadic eye movements to targets.

Application of pharmacological agents could also alter the properties of the models. Agonists acting at all inhibitory synapses would silence the oculocentric maps of both models, since all units receive such synapses from inhibitory interneurons. The activity of the intermediate units would become irrelevant in the vector subtraction model. An antagonist acting selectively at the synapses made by the eye position units but not at the synapses made by the inhibitory interneurons would have the same effect as eliminating the eye position signal. Antagonists acting at the synapses made by the inhibitory interneurons as well would cause the oculocentric units in the vector subtraction model to lose much of the head-centered spatial selectivity that they would retain if only the eye position signal were lost. The loss of the influence of the inhibitory interneurons would permit the units located centrally in the map to be active for high levels of activity in the intermediate units. The receptive fields would be head-centered but unbounded on the periphery. In the dendrite model, loss of both the inhibitory interneurons and the eye position units would allow the oculocentric units to respond to targets from a much broader region of cranioocentric space. Since the mapping from the cranioocentric map onto the oculocentric map is not complete, however, the oculocentric units would exhibit broad spatial tuning in head-centered coordinates.

It is interesting to speculate on the effects of anesthesia with respect to the model since so much of the information on the auditory system comes from anesthetized animals. Anesthesia may directly affect the circuitry responsible for computing the location of an auditory target in oculocentric coordinates, by enhancing or reducing the efficacy of the synapses employing specific neurotransmitters in some of the ways discussed above. It may also indirectly affect the signals entering this circuitry through actions elsewhere. Barbiturate anesthesia has been shown to affect the responses of neurons in the IC for example (Kuwada et al. 1989).

Anesthesia clearly affects the computation of the coordinate transformation by altering the eye position signal. The eye position signal probably arises from a corollary discharge of a motor command to move the eye. It is conceivable that some form of commands for eye movements might still occur in the anesthetized or paralyzed animal although the eyes do not actually move. Thus the eye position signal might erroneously vary across time despite the stability of the eyes in the orbit. In that case, oculocentric auditory units in both models which would otherwise be sensitive to eye position would appear to have highly variable responses under anesthesia. These responses might be mistaken for a head-centered receptive field due to the limitations on the projections from the cranioocentric map to the oculocentric map in the dendrite model.

3.5 Applications to the visual system

The issue of the coding of visual information with respect to the head has gained attention recently. A
head-centered code of visual signals is thought to be a way in which perception of a stable visual environment across time might be achieved despite intervening eye movements. Such a coordinate frame might facilitate binding of information of different sensory modalities arising from the same object. A hint of head-centered coding has been found in Area 7a of parietal cortex, where neurons have retinal receptive fields with eye position gain fields (Andersen et al. 1985; Andersen and Zipser 1988; Zipser and Andersen 1988).

Since visual signals arise in eye-centered (retinotopic) coordinates, the computation to transform the signal into head-centered coordinates is the exact opposite of that required for the auditory head- to eye-centered transformation: the addition of a signal of the position of the eyes with respect to the head to the retinal location of the target. Only slight modifications are needed to convert the auditory models into visual models for executing the reverse transformation.

The vector subtraction model becomes the vector addition model (Fig. 6). The eye position units form excitatory synapses on the intermediate units. The pattern of synaptic weights in the output layer is rescaled according to the higher range of activity in the intermediate units. The dimensions of the input and output maps stay the same. Now the input map represents \( \pm 20^\circ \) on the retina for one eye position at a time, and the output map represents the \( \pm 20^\circ \) of visual field for eye positions \( \pm 20^\circ \) in the oculomotor range. One change remains: a new source of inhibition is necessary to ensure that the output map is not active unless a visual target is actually present. This can be accomplished by a second projection from the retinal map to the intermediate units via an inhibitory interneuron. The projection onto the interneuron is inhibitory. Accordingly, when a target is present, the inhibitory interneuron will be turned off, releasing the intermediate unit and allowing it to fire.

This model is very similar to a neural network model proposed by Zipser and Andersen 1988; Andersen and Zipser 1988; Goodman and Andersen 1989. Their three layer network consists of place-coded retinal input, a layer of hidden units, and an output layer encoding target location with respect to the head in the form of either a rate-code or a place-code. The synaptic weights of the hidden units were trained using backpropagation. A detailed comparison with the current models serves to highlight the essential features of each.

Zipser and Andersen prefer the version of their model which uses rate-coded output because the hidden units more closely resemble parietal cortex neurons in that they have retinal receptive fields with planar eye position gain fields. Few of the hidden units of the model with place-coded output are of this type. Their rate-coded model corresponds most closely with the first two layers of the vector subtraction model, as modified above. The input layers of both represent visual signals in an eye-centered place-code. The intermediate stage of the vector addition model contains a rate-code of the visual target with respect to the head like the output layer of the Zipser and Andersen model. From the similarity it can be seen that the Zipser and Andersen model would also fail to represent multiple visual stimuli. In fact, their model was trained and tested with single targets only. As with the auditory vector subtraction model, if the model is correct it must lie downstream from some attentional filter or target selection mechanism that chooses a single target to be transformed from retinal to head-centered coordinates. This model is not appropriate for transforming an entire visual scene from retinal to head-centered coordinates, and thus does not serve the arguments for a head-centered frame of visual reference to account for perceptual stability as the eyes move.

The dendrite model can also be modified to transform visual signals from retinal to head-centered coordinates. The architecture need not be changed at all. The input layer now encodes retinal location. The weights of the synapses on the dendrites will be different, but the computation performed by the local circuitry will be identical: selecting combinations of retinal locations and eye positions. The relative dimensions of the input and output maps will be the same as for the auditory map, with the input map being smaller than the output. Zipser and Andersen's model may have suffered from edge effects since their input and output maps were the same size and therefore mismatched.

Fig. 6. Vector addition model. Additional inhibitory interneurons have been added at the intermediate units.
The dendrite model can transform an entire visual scene from retinal to head-centered coordinates, but it does not contain hidden units resembling parietal cortex. Even if the convergence and divergence of the projection from retinal units to head-centered units occurred in more than one step, the intermediate stages would differ significantly from parietal neurons. If intermediate units received input only from a subset of retinal input units, they would appear to have head-centered receptive fields gated by retinal position, not retinal receptive fields with eye position gain fields. It would be easy to produce parietal neurons in a model by simply adding rate-coded eye position signals to units with retinal receptive fields, but such units are clearly not a necessary component of the transformation from retinal to head-centered coordinates.

These models show that transforming between head- and eye-centered coordinate frames need not be a complicated process. The differences between them highlight the costs and benefits associated with the different neural algorithms. Use of a rate-coded intermediate stage in the vector subtraction model permits a minimal investment of neurons to execute the mathematics of the coordinate transformation for single targets. The dendrite model suggests a reasonable scheme for transforming a larger amount of information by making use of independent dendritic processing, while also keeping the number of neurons at a minimum. These are simple and neurally plausible ways in which the brain might accomplish the task of allowing sensory signals originally encoded in different frames of reference to share a common motor pathway.

4 Appendix: general solution for the dendrite model

The input to an oculocentric unit via any dendrite is the sum of the product of the synaptic weights and the firing rates of the presynaptic units. Synapses are made by one cranio-centric auditory unit, one horizontal eye position unit, one vertical eye position unit, one horizontal inhibitory interneuron and one vertical eye position unit. (See Table 2 for the synaptic weights and thresholds.) The unit’s response is binary: 0 if the input is less than threshold, and 1 if the input is greater than or equal to threshold = 0.

Let the auditory target be at position \( x(a_n), y(a_n) \) and the eyes at \( x(e_h), y(e_h) \). Then,

\[ x(a_n) = x(a_n) - x(e_h) \]
\[ y(a_n) = y(a_n) - y(e_h) \]

and,

\[ x(e_h) = x(a_n) - x(a_n) \]
\[ y(e_h) = y(a_n) - y(a_n) \]

We will consider the oculocentric unit whose receptive field is at \( (x(a_n), y(a_n)) \), where both \( x(a_n) \) and \( y(a_n) \) > 0. We will analyze the activity at the dendrite receiving input from the cranio-centric unit \( (x(a_n), y(a_n)) \).

The input to the horizontal inhibitory interneuron of the dendrite is

\[ -(k + x(e_h)) \]
\[ k + x(a_n) - x(a_n) \]

Substitution gives

\[ -(k + x(e_h)) \]
\[ k + x(e_h) - 1 \]

This is less than 0, so unit does not fire.

VI’s input is similar, and also does not fire. Therefore the input to the dendrite itself is

\[ -(k + y(e_h)) \]
\[ (k + y(a_n) - y(a_n)) \]

Substitution gives

\[ -(k + y(e_h)) \]
\[ k + y(e_h) - 0 \]

which is equal to threshold, so the unit fires.

If the value of either \( x(e_h) \) or \( y(e_h) \) were greater, then the dendrite would be inhibited by the direct input from the eye position units. If these values were smaller, the inhibitory interneurons would become active and shut down the dendrite. If the target were at a different position in head-centered space, then a different cranio-centric unit would be active, and a different dendrite on this oculocentric unit would receive excitation, and a different eye position would be necessary to allow this excitation to cause the cell to fire.

Acknowledgements. The authors thank LH Finkel, P Sadka, PW Glimcher, MJ Nichols, SJ Nowlan, TR Stanford and VM Caramitaro for their assistance. This research was supported by National Institutes of Health grant R37 EY01189 to DLS, and by National Science Foundation and National Defense Science and Engineering graduate fellowships to JMG.

References


Schiller PH, True SD, Conway JL (1979) Paired stimulation of the frontal eye fields and the superior colliculus of the rhesus monkey. Brain Res 179:162–164


