

A model of target selection based on goal-dependent modulation

Emilio Salinas¹

*Department of Neurobiology and Anatomy, Wake Forest University School of Medicine,
Winston-Salem, NC 27157, USA*

Abstract

When contemplating a crowded scene, such as a dinner table, gaze can be directed toward a large number of potential targets — a glass, a fork, a salt shaker — any one of which can be specifically sought and viewed. Their presence is signalled by the evoked responses of visual neurons with appropriate selectivities, but if the item designated as the current target changes from moment to moment, how can these sensory responses be used to direct eye movements? Additional activity encoding the current goal must be taken into account. Here I show that if this contextual activity is combined nonlinearly with the sensory information, then the proper oculomotor command needed to foveate the current target can be easily read out by downstream neurons. This mechanism is illustrated with a model network that solves a target-selection task. The model proposes that, during visual search, sensory responses should be nonlinearly modulated by the identity of the goal. This modulation effectively changes the functional connectivity of a network, allowing it to locate different targets under different circumstances.

Key words: Gain fields, Basis functions, Arbitrary visuomotor remapping, Neural network, Saccade.

1 Introduction

Visual search tasks may be complex even when the object recognition step is obviated or considered separately. This is because a sought object may appear at multiple locations and because the object considered as the target may change from one moment to the next, depending on ongoing needs,

¹ Supported by NINDS grant NS044894-02.

E-mail address: esalinas@wfubmc.edu.

motivations and other circumstances. How is all this contextual or goal-dependent information used to single out a small fraction of the sensory landscape and influence motor behavior?

Neurons often have a variable gain, which allows them to integrate information from various modalities or sources [13,12]. Such 'gain-modulated' (GM) neurons have a sensory receptive field, but in addition, their overall responsiveness depends on other modulatory quantities that do not affect their selectivity. A well known example is that of parietal neurons, whose visual receptive fields are gain modulated by the direction of gaze [1,3]. Many other cases of modulation by internal or proprioceptive signals are known [13,12]. Such modulation is thought to be the basis for performing coordinate transformations; for instance, from eye-centered to body-centered representations [14,10,7]. More generally, however, nonlinear modulation can be used to compute any function of the encoded quantities [6,7,11,9].

Here, through computer simulations of a model network, I show that this mechanism can be successfully applied to the problem of target selection. The key is that the stored information about the current goal must have a nonlinear effect on the visually-triggered responses of some neurons.

2 A Simple Target-Selection Task

The behavioral task performed by the model is illustrated in Figs. 1a, b. Two circles, an open and a filled one, are presented in each trial while the subject fixates. The task is either to shift gaze from the fixation point to a target circle (in go trials), or to maintain fixation (in no-go trials). A separate cue at the beginning of each trial indicates which of three conditions apply: the filled circle is the target (Fig. 1a), the open circle is the target (Fig. 1b), or the trial is no-go. The circles appear within a fixed range along the horizontal direction and their positions vary across trials. Each circle is thus associated with multiple eye movements, and conversely, any particular saccade vector points to the filled circle in some trials and to the open one in others.

3 A Two-Layer Network Model

The network that simulates the task consists of a layer of GM neurons that drive a second layer of output or motor neurons through a set of connections. The output population is meant to encode the location of a target for an impending movement. In every trial of the task, first, the GM neurons

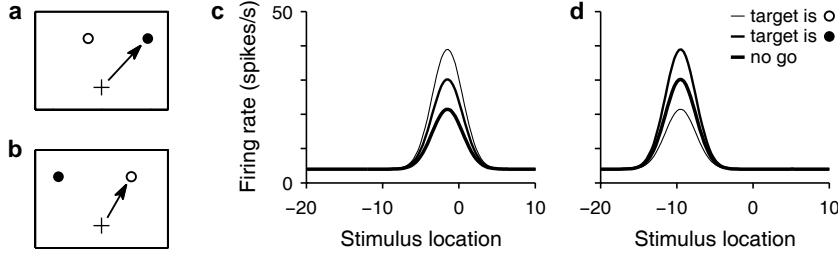


Fig. 1. Target-selection task and gain-modulated (GM) responses. (a) Schematic of the task when the correct response is an eye movement (arrow) from the fixation point (cross) to the filled circle, which is the target. (b) Schematic of the task when the open circle is the target. Circle positions vary horizontally from trial to trial. Target identity is revealed by a brief cue (not shown) starting each trial. Alternatively, the cue may indicate that the trial is no-go, in which case fixation should be maintained. (c) Response of a model GM neuron as a function of the location of its preferred stimulus (the open circle). Its response amplitude (gain) varies with the goal; the maximum gain is obtained when the open circle is the target. (d) As in (c) but for another GM neuron. This unit also prefers the open circle in its receptive field, but the maximum gain is obtained when the filled circle is the target. Note different preferred locations.

are activated by the sensory input and are modulated by the cue indicating the goal. Then, in go trials, a movement is generated by the output neurons; for this, their profile of activity should have a single peak at the location of the movement to be made. In the no-go condition, however, all output responses should stay at their low, baseline level, so the profile should be flat.

The visual response of GM neuron j depends on how strongly it prefers one stimulus (open or filled) versus the other, and on where those stimuli are relative to the neuron's preferred location x_j . In Figs. 1c, d, the preferred locations are the points on the x axis that correspond to the peaks of the curves. For simplicity, the activities evoked by the two simultaneously presented stimuli are combined linearly. In addition, each visual response is multiplied by a gain factor that depends on the goal (target is open, target is filled, or no-go). The goal is indicated by y , which may take 3 values, 1, 2 or 3. Putting together all of these dependencies, the mean firing rate r_j of GM unit j is written as

$$r_j = r_{max} \left(A_j^f f(x^f - x_j) + A_j^o f(x^o - x_j) \right) \left(1 - D + D g_j(y) \right) + B, \quad (1)$$

where B is a baseline rate equal to 4 spikes/s, $r_{max} = 35$ spikes/s, and D is the modulation depth; here, $D = 0.5$, which produces a maximum goal-dependent suppression of 50%. The function f is a Gaussian (see Figs. 1c, d). It appears twice, once for the filled circle, which is located at x^f and once

for the open circle, which is at x^o . The terms A_j^f and A_j^o reflect the preference for filled versus open circles. The neurons in Fig. 1c, d do not respond to the filled circle, so they have $A_j^o = 1$ and $A_j^f = 0$. For each cell in the full network, however, one of these terms (chosen randomly) is equal to 1 and the other is equal to a random number between 1 and 0.2.

The functions $g_j(y)$ control the gain. They take 3 values, one per goal or task condition. These are $g_j(y) = \{1, 0.5, 0\}$. Crucially, they are assigned randomly to each of the 3 goals, with a new random permutation for each GM unit. Thus, the idea is to generate an array of GM cells with as many combinations as possible of preferred stimulus, preferred location and preferred goal. This is essential [7,8,4,9], and gives rise to seemingly paradoxical responses, like the one in Fig. 1 d: the maximum response occurs when the open circle is in the receptive field but the filled one is the target.

The firing rate of each output neuron in the model is determined by a weighted sum of GM rates,

$$R_i = \sum_j w_{ij} r_j, \quad (2)$$

where R_i is the firing rate of output unit i and w_{ij} represents the synaptic connection from GM neuron j to output neuron i . This expression is used when the GM neurons drive the output neurons. However, there is also an intended or desired response F for the output neurons, which is used only when setting the connections. In no-go trials ($y = 3$), all neurons should keep firing at the baseline rate, so $F = B$. In go trials, the profile of activity should be a Gaussian centered at the location of the target, so $F = F(x^o - c_i)$ when the open circle is the target ($y = 1$) and $F = F(x^f - c_i)$ when the filled one is the target ($y = 2$). In these expressions, F is a Gaussian function and c_i is the preferred movement location of output neuron i . The center of mass of the output activity is interpreted as the encoded target location; this is

$$T_{out} = \frac{\sum_i (R_i - B)^2 c_i}{\sum_k (R_k - B)^2}. \quad (3)$$

The connections to any given output neuron are chosen so that they minimize the average squared difference between intended and driven output responses. The method to find the optimal weight values is the same as in [9]. This algorithm is run only once, before testing the model; afterward the synaptic weights are not modified any further.

4 Performance of the Model

In each trial of the task, the model is tested as follows: (1) stimulus locations x^f and x^o and goal y are chosen, (2) all GM responses are generated (Eq. 1), (3) the driven output firing rates are calculated (Eq. 2), and (4) the encoded movement location is read out (Eq. 3) and compared to the location of the target (in go trials only). Note that, to evaluate the robustness of the model, noise is added to all GM responses in every trial.

Fig. 2a illustrates the behavior of the full model in 6 single trials. The plots show the profile of activity of the 25 output neurons in the network driven by 2000 GM units. The positions of the stimuli are indicated in each case.

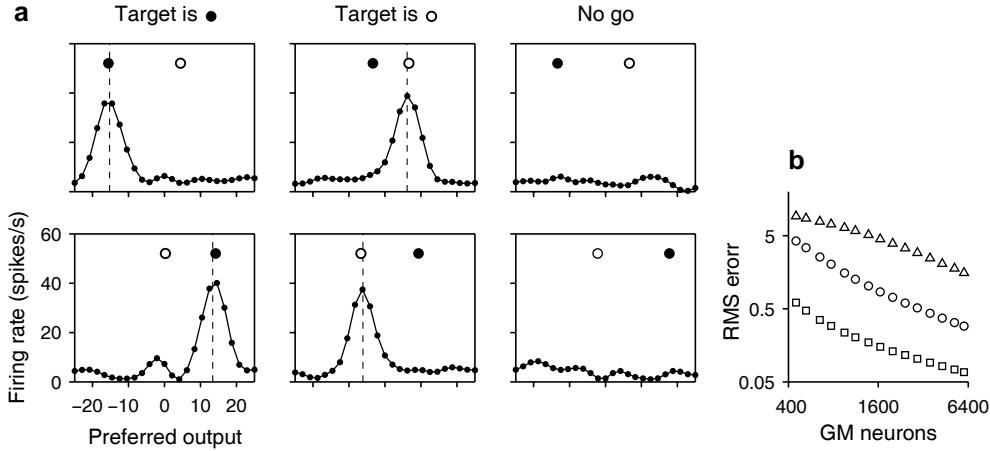


Fig. 2. Performance of the network. (a) Six single trials. The network had 2000 GM and 25 output neurons. In each trial, the two stimuli could be placed anywhere in the range $[-20, +20]$, with a separation between 10 and 20 units. Plots show output responses ordered by their preferred movement locations c_i . Stimulus locations are as indicated by the open and filled circles. Each column corresponds to a different goal, as indicated. Dashed lines mark the center of mass of the output activity. (b) RMS difference between target location and movement location (center of mass) as a function of network size. Each point is an average over multiple stimulus locations, target conditions and trials. The three curves are for different noise levels: $\alpha = 0.25$ (squares), $\alpha = 1$ (circles), and $\alpha = 4$ (triangles), where the variance of the noise for each GM response r_j was equal to αr_j . In panel (a), $\alpha = 1$.

The model performs accurately because, in go trials, both the peak response and the center of mass (dashed line) are near the target location, and this is true for both circles and for all positions. Furthermore, in no-go trials the profile of activity is approximately flat.

The activity profiles in Fig. 2a have secondary peaks and irregularities. These are due to noise; with zero noise they disappear completely, but oth-

erwise they tend to shift the center of mass away from the correct target location. The misalignment caused by the fluctuations can be reduced by increasing the number of GM neurons, N . This is shown in Fig. 2b, which plots the root-mean-square (RMS) difference between movement and target locations as a function of N . The decrease in RMS error indicates that noise is indeed what limits the accuracy of the network.

5 Discussion

The model presented here generates a movement toward a target that may appear in any position. If only one condition is considered, say, when the target is the open circle, the network’s function becomes trivial: a one-to-one map between stimulus and movement locations. What is interesting is that the same neurons, driven by the same stimuli, can implement an entirely different map when the filled circle is the designated target. Furthermore, similar patterns of GM activity give rise to a flat or null response in the no-go condition. Thus, three maps are overlaid on the same set of connections, and it is the pattern of goal-dependent modulations that expresses one or another map. Note that the changes in GM activity needed to achieve this were relatively subtle: the maximum decrease in a GM rate between two conditions was 50% (Eq. 1),

This model uses the same mechanism proposed for performing coordinate transformations [14,10,7] and generating visual responses that appear to be object-centered [4]. However, in this case the modulatory quantity is neither proprioceptive (like eye position [1,3]) nor present at the same time as the stimuli. Such contextual modulation may, in general, serve to select among various visuomotor maps [9], and as shown here, may also allow a motor network to read out the location of one of many potential targets.

This model may be extended in several ways. It should be able to deal with variable numbers of distractors and with more than one feature, for instance. The effects of GM receptive field nonlinearities and of lateral interactions between neurons may also be important. But the essential predictions depend on the idea of generating a set of neuronal responses that behave as basis functions [6–8,2], and thus would not change: during visual search, some neurons should mix sensory and goal-dependent signals, the mixing should be nonlinear, and a linear combination of responses should produce a highly accurate estimate of target location. Consistent with this, there is evidence indicating that visual responses in area V4 are modulated by the identity of the target [5], but further work is needed to test these predictions rigorously.

References

- [1] R.A. Andersen, G.K. Essick, R.M. Siegel, Encoding of spatial location by posterior parietal neurons. *Science* 230 (1985) 450–458.
- [2] S. Ben Hamed, W. Page, C. Duffy, A. Pouget, MSTd neuronal basis functions for the population encoding of heading direction. *J. Neurophysiol.* 90 (2003) 549–558.
- [3] P.R. Brotchie, R.A. Andersen, L.H. Snyder, S.J. Goodman, Head position signals used by parietal neurons to encode locations of visual stimuli. *Nature* 375 (1995) 232–235.
- [4] S. Deneve, A. Pouget, Basis functions for object-centered representations. *Neuron* 37 (2003) 347–359.
- [5] J.A. Mazer, J.L. Gallant, Goal-related activity in V4 during free viewing visual search: evidence for a ventral stream visual salience map. *Neuron* 40 (2003) 1241–1250.
- [6] R. Poggio, A theory of how the brain might work. *Cold Spring Harbor Symp. Quant. Biol.* 55 (1990) 899–910.
- [7] A. Pouget, T.J. Sejnowski, Spatial transformations in the parietal cortex using basis functions. *J. Cog. Neurosci.* 9 (1997) 222–237.
- [8] A. Pouget, L.H. Snyder, Computational approaches to sensorimotor transformations. *Nat. Neurosci.* 3 (Suppl) (2000) 1192–1198.
- [9] E. Salinas, Fast remapping of sensory stimuli onto motor actions on the basis of contextual modulation. *J. Neurosci.* 24 (2004) 1113–1118.
- [10] E. Salinas, L.F. Abbott, Transfer of coded information from sensory to motor networks. *J. Neurosci.* 15 (1995) 6461–6474.
- [11] E. Salinas, L.F. Abbott, Invariant visual responses from attentional gain fields. *J. Neurophysiol.* 77 (1997) 3267–3272.
- [12] E. Salinas, T.J. Sejnowski, Gain modulation in the central nervous system: where behavior, neurophysiology and computation meet. *Neuroscientist* 7 (2001) 430–440.
- [13] E. Salinas, P. Thier, Gain modulation: a major computational principle of the central nervous system. *Neuron* 27 (2000) 15–21.
- [14] D. Zipser, R.A. Andersen, A back-propagation programmed network that simulates response properties of a subset of posterior parietal neurons. *Nature* 331 (1988) 679–684.