Neuron **61**:168-185, 2009

The Normalization Model of Attention

Supplementary Material

John H. Reynolds[1] and David J. Heeger[2]

[1] Salk Institute for Biological Studies
reynolds@salk.edu

[2] Department of Psychology and Center for Neural Science, New York University
david.heeger@nyu.edu

Supported by NEI grant R01-EY016161 (JHR) and NIMH grant R01-MH69880 (DJH).

Notation

Capital letters are used to denote neural images, or the collected responses of populations of simulated neurons:

$R(x,\theta)$ is the output firing rates of a population of neurons with x denoting the receptive field centers and θ denoting the orientation (or direction) preferences.

$E(x,\theta)$ is the underlying stimulus drive for the entire population.

$S(x,\theta)$ is the corresponding suppressive drive for the entire population.

$A(x,\theta)$ is the attention field, across the entire population.

$R(x,\theta; c,\phi)$ represents the output firing rates as a function of stimulus contrast (c) and orientation/direction (ϕ) for a neuron with receptive field center x and orientation/direction preference θ.

$E(x,\theta; c,\phi)$ represents the stimulus drive for the entire population of neurons in response to a stimulus with contrast c and orientation ϕ.

$A(x,\theta; \phi)$ represents the attention field for the entire population of neurons when attending to a particular orientation/direction (ϕ).

Lower case letters are used to characterize an individual simulated neuron in the population:

$r(c)$ is the contrast-response function with c denoting stimulus contrast.

$r(\phi)$ is the orientation/direction tuning curve with ϕ denoting stimulus orientation (or direction).

$e(\phi)$ is the orientation/direction tuning of the underlying stimulation field, i.e., the stimulus drive for a particular neuron in the population in response to various different stimulus orientations/directions (ϕ).

$a(\phi)$ is the orientation/direction dependence of the attention field, i.e., the attentional gain of a particular neuron when attending to various different orientations/directions (ϕ).

Derivation of model predictions corresponding to Figure 4C

A model neuron with receptive field center x and orientation preference θ is presented with two stimuli in its receptive field, one moving in the preferred direction and the other moving in the non-preferrred (opposite) direction. The contrast of the preferred direction stimulus (c_p) varied across trials, whereas the contrast of the non-preferred stimulus (c_n) was fixed at 10%. The responses of the model neuron, when attending to the stimulus in the opposite hemifield (i.e., ignoring the stimulus in the receptive field), can be expressed (see equation [4]) as:

$$r_i(c_p) = \alpha\, c_p / (c_p + \beta\, c_n + \sigma), \qquad [S1]$$

where $0 < \beta < 1$ is a scale factor on the suppressive drive from the non-preferred direction of motion (analogous to the suppression from surrounding locations in equation [4]). When attending the non-preferred stimulus within the receptive field:

$$r_n(c_p) = \alpha\, c_p / (c_p + \gamma\, \beta\, c_n + \sigma), \qquad [S2]$$

where $\gamma > 1$ is the peak gain of the attention field. The attentional gain is multiplied only by c_n in the denominator because only the non-preferred stimulus is attended. Because c_n was fixed in this experiment, equations [S1] and [S2] can be rewritten:

$$r_i(c_p) = \alpha\, c_p / (c_p + \sigma') \qquad [S3]$$

$$r_n(c_p) = \alpha\, c_p / (c_p + \sigma''), \qquad [S4]$$

where $\sigma' = (\beta\, c_n + \sigma)$ and $\sigma'' = (\gamma\, \beta\, c_n + \sigma)$ are constants that represent the neuron's contrast gain. Attending the non-preferred stimulus in the receptive field decreases the neuron's contrast gain by a factor of σ' / σ'', that is, a rightward shift when plotting the responses versus the logarithm of contrast (Figure 4C).

Derivation of model predictions corresponding to Figure 4E

The same model neuron is presented with the same two stimuli within its receptive field, except that the contrasts of both stimuli are covaried ($c_n = c_p$). The responses, when attending to the preferred (r_p) and non-preferred (r_n) stimuli, can be expressed as:

$$r_p(c) = \alpha\, \gamma\, c / (\gamma\, c + \beta\, c + \sigma), \qquad [S5]$$

$$r_n(c) = \alpha\, c / (c + \gamma\, \beta\, c + \sigma). \qquad [S6]$$

For low contrasts ($c << \sigma$), equations [S5] and [S6] can be approximated as:

$$r_p(c) \approx \alpha\, \gamma\, c / \sigma \qquad [S7]$$

$$r_n(c) \approx \alpha\, c / \sigma, \qquad [S8]$$

such that the responses are predicted to be larger by a factor of γ when attending the preferred stimulus. For high contrasts ($c >> \sigma$), the responses can be approximated as:

$$r_p(c) \approx \alpha\, \gamma / (\gamma + \beta) \qquad [S9]$$

$$r_n(c) \approx \alpha / (1 + \gamma\, \beta), \qquad [S10]$$

which is independent of c because the responses saturate at high contrasts. Predicted responses are again larger when attending the preferred stimulus, $r_p > r_n$. For this experiment, therefore, the model predicts predominantly a change in response gain rather than a change in contrast gain (Figure 4E).

Derivation of model predictions corresponding to Figure 5C

A model neuron with receptive field center x and orientation preference θ is presented with stimuli centered in its receptive field, with various contrasts c and various stimulus orientations ϕ. The responses of this neuron can be expressed as:

$$R(x,\theta; c,\phi) = \left| A(x,\theta)\, E(x,\theta; c,\phi) / [s(x,\theta) * A(x,\theta)\, E(x,\theta; c,\phi) + \sigma] \right|_T , \qquad [S11]$$

where ϕ represents the stimulus orientation. In this simulation, attentional modulation is independent of orientation, i.e., $A(x,\theta) = A(x)$. Assuming that the suppressive field $s(x,\theta)$ pools equally across all orientations, the denominator will not depend on ϕ, only on contrast. The stimulus drive can be expressed as $E(x,\theta; c,\phi) = \alpha\, c\, e(\phi)$, proportional to contrast times a contrast-independent orientation tuning curve $e(\phi) = E(x,\theta; 1,\phi)$. Hence, we find that the output firing rates can be approximated as a separable function of contrast and orientation:

$$r(c,\phi) = \alpha\, g(c)\, e(\phi) \qquad [S12]$$

$$g(c) = [c / (c + \sigma/\gamma)].$$

This approximation would be exact if the attention field were large (as in equation [7]) and the suppressive field were pooled equally across all orientations. The attentional gain, γ, is greater than 1 when attending to the receptive field, but equal to 1 when attending to the stimulus in the opposite hemifield. Attending to the receptive field ($\gamma > 1$), changes the contrast gain of g(c) but has no effect on e(ϕ), such that the predicted tuning curve for attended stimuli is a multiplicatively scaled copy of the tuning curve for unattended stimuli (Figure 5C).

Derivation of model predictions corresponding to Figure 6C

A model neuron with receptive field center x and direction preference θ is presented with stimuli of various motion directions centered in its receptive field. Spatial attention is always directed away from the stimulus in the receptive field, but feature-based attention is matched to the receptive field stimulus on half the trials. The attention field in this simulation is, therefore, selective for motion direction. Analogous to the derivation of equation [S12], the direction tuning, $r(\phi) = R(x,\theta;1,\phi)$, of the simulated neuron can be approximated as:

$$R(x,\theta;1,\phi) = \lfloor A(x,\theta;\phi)\, E(x,\theta;1,\phi) \,/\, [s(x,\theta) * A(x,\theta;\phi)\, E(x,\theta;1,\phi) + \sigma] \rfloor_T \quad [S13]$$

$$r(\phi) = a(\phi)\, e(\phi)\, [\alpha / (\gamma k + \sigma)],$$

$$a(\phi) = A(x,\theta;\phi),$$

$$e(\phi) = E(x,\theta;1,\phi),$$

where α is again a proportionality constant, $k > 0$ is a constant, ϕ now represents the stimulus direction (rather than orientation), e(ϕ) is the direction tuning of the underlying stimulation field, and a(ϕ) is the motion-direction dependence of the attention field. The approximation in equation [S13] would be exact if the suppressive field were constant across all motion directions such that the term in the denominator, $s(x,\theta) * A(x,\theta;\phi)\, E(x,\theta;1,\phi)$, is equal to the γk for all motion directions. With attention directed to the fixation point, $a(\phi) = 1$ (and $\gamma=1$), the direction tuning is simply proportional to e(ϕ). With feature-based attention, the tuning is predicted to be proportional to the product a(ϕ) e(ϕ), which is narrower because the attention field is selective for motion direction (Figure 6C).

Derivation of model predictions corresponding to Figure 7C

Two stimuli are presented within a model neuron's receptive field, one of which moves in the non-preferred direction, while the other varies in motion direction (ϕ). Attention is directed either to the fixation point, or to one of the two stimuli in the receptive field. The responses of the model neuron can again be approximated by assuming that the suppressive field is constant across all motion directions. Under this approximation:

$$r_f(\phi) = \alpha\, e(\phi) / (k + \sigma), \quad [S14]$$

$$r_n(\phi) = \alpha\, e(\phi) / (\gamma k + \sigma), \quad [S15]$$

$$r_v(\phi) = \alpha\, a(\phi)\, e(\phi) / (\gamma k + \sigma), \quad [S16]$$

where $r_f(\phi)$ is the direction tuning when attending fixation, $r_n(\phi)$ is the tuning when attending the non-preferred motion direction, and $r_v(\phi)$ is the tuning when attending the variable motion direction. The attentional gain, γ, appears in the denominator when attending one of the moving stimuli in the receptive field, regardless of its direction, because the suppressive field is assumed to be constant across all motion directions. The model makes three predictions. First, attending the non-preferred stimulus should reduce the neuronal responses: $r_n < r_f$. Second, responses should be larger when attending the stimulus with variable motion direction: $r_v > r_f$ for motion directions such that $a(\phi) \approx \gamma$ because $\gamma / (\gamma k + \sigma) > 1 / (k + \sigma)$; and $r_v \geq r_n$ for all motion directions because $a(\phi) \geq 1$. Third, the model also predicts a sharpening of the tuning curve when attending the variable motion direction (note a(ϕ) in the numerator of equation [S16] like that in the second line of equation [S13]). The degree of sharpening depends on the width of the attention field, a(ϕ). An attention field with flat feature-tuning (equal for all feature values) would have resulted in a scaling of responses with no sharpening of tuning. The attention field in this simulation was relatively broad, but not completely flat, resulting in only a modest sharpening of tuning in Figure 7C.