Appendix S1

Models of Attentive Neural Modulation – Computational Details

The descriptions and equations included below are meant mainly to illustrate the wide variety of solutions proposed, and to highlight the way attentional modulation is implemented, rather than being exhaustive descriptions of the models. For full details, complete sets of equations and biological justification, the reader is referred to the original sources.

The Biased Competition model [19] has been proposed as a demonstration of the biased competition theory. The temporal evolution of a neuron’s firing rate is given by:



where  is the excitatory input for two stimuli within the same receptive field, and the inhibitory input for two stimuli within the same receptive field. The *w*'s are positive and negative synaptic weights. The equilibrium response is described by  , where *B* is the maximum response and *A* is a decay constant. Attention is assumed to increase the strength of the signal coming from the inputs activated by the attended stimulus, implemented by increasing the associated synaptic weights.

The Neurodynamical model [21] presents a large-scale implementation of biased competition that consists of several interconnected network modules simulating different areas of the dorsal and ventral path of the visual cortex. Each module consists of a population of cortical neurons arranged in excitatory and inhibitory pools. The temporal evolution of the system is described within the framework of a mean-field approximation, i.e. an ensemble average of the neural population is calculated in order to obtain the corresponding activity. For example, the current activity of the excitatory pools in the posterior parietal (PP) module is given by:



where  is an external attentional spatial-specific top-down bias,  is a diffuse spontaneous background input,  is Gaussian noise. The intermodular attentional biasing  through the connections with the pools in the module V1 is:



and the activity current of the common PP inhibitory pool evolves according to:



Similar equations govern the dynamics of the other models.

The Feedback Model of Visual Attention [20] improves on the Biased Competition Model by providing a biologically-justified mechanism and microcircuitry for input modulation. The key observation that drives the model is that feedforward connections seem to be primarily made in basal dendrites, while feedback connections preferentially target apical dendrites, thus appearing to have functionally different roles. The activations of the apical and basal dendrites are given by:

 and 

where *j* and *k* are indices used to localize the neuron in the network, *ma* and *mb* are the total number of apical and basal synapses, respectively, *vijk* and *wijk* are synaptic weights associated with the input from neuron *i* (apical and basal, respectively), *x* is the output of neuron i. The apical inputs  originate from higher cortical regions or are top-down signals from outside the model. The basal inputs are the outputs of neurons in lower cortical regions, after pre-integration lateral inhibition:



where αt scales lateral inhibition,  is the output of neuron p in region k at the previous time step, and  the positive half-rectified value of . The Feedback Model of Visual Attention proposes that feedback activations multiplicatively modulate the total feedforward activation:

 .

The reentry hypothesis [22] models top-down modulation as a gain control mechanism on the input feedforward signal, increasing activations that match top-down predictions. Considering two interconnected areas I and II, the generic formulation of the modulation of the input signal  to area II combines the filtered feedforward input  from area I with the summed top-down signal  where *γ∈{L,F}* is the origin of the attentional signal (*L* for location, *F* for feature):



The nonlinear pooling function *f* defines the influence of the filtered afferents *F* on cell *k*. This generic equation is adapted to the specific connectivity of each area.

In the computational model of feature similarity gain proposed by [25], the neural response is described as a divisive contrast normalization process between the sum of squared linear responses to each stimulus component and the sum of squared contrasts plus a semi-saturation constant *σ* by:



where each component that contributes to the activation has its own feature *xi* and contrast *ci*. The normalized firing rate *H* is modulated by a gain factor *G(y)* that has a tuning function similar to the stimulus-driven tuning function of that neuron, where *y* is the attended feature, resulting in a modulated response given by :

 .

The parameter *δ* is the inherent baseline-firing rate of the neuron. The gain factor *G(y)* is greater than 1.0 for a preferred feature, and lower than 1.0 otherwise. The gain factor is a purely feature-based effect, and is independent of the spatial focus of attention and the properties of the visual stimulus.

Attentional modulation in saliency models is investigated in the context of object recognition [27]. A focus of attention spatial region is determined by means of center-surround competition along different feature dimensions and scales. A modulation mask is obtained by rescaling, smoothing and normalizing the focus of attention to the resolution of the layer where attention is to be applied (corresponding to visual areas V1 or V4)



If *S(x,y)* is the neural activity at position *(x,y)*, the modulated activity is computed according to



where 0≤μ≤1 is the strength of the attentional modulation. Modulation factors of 0.2-0.4 are shown to be sufficient for maximizing recognition performance in a multi-stimulus experimental setup.

The Normalization Model of Attention [30] combines neural selectivity (termed “stimulus drive”) with an external “attention field” and a “suppressive field”, that pools activations corresponding to non-preferred stimulus and unattended locations, which is used as in normalization. The resulting firing rates are defined as:



where

.

*E(x,θ)* is the stimulus drive at location *x* for orientation *θ. A(x,θ*) is the attentional field, *S(x,θ)* is the suppressive drive. *σ* is a constant that determines the neuron's contrast gain. |...|*T* specifies rectification with respect to threshold *T.* ✴ denotes convolution, *s(x,θ)* gives the extent of pooling. Stimulus contrast is also included in the equations and is not shown here; see [30] for further details. *A(x,q*) *=1* everywhere except for the positions and features that are to be attended, where *A(x,q) >1.*

The Normalization Model of Attentional Modulation [31] proposes that the primary effect of attention is to modulate the strength of normalization mechanisms. With two stimuli present in the RF, the neuron’s firing rate is given by:



where *Ni* is the normalization term for each stimulus, *Ii* is the direct input driven by each stimulus, and *u* is a power term that enables the modeling of different nonlinear summation regimens. Each normalization term depends on the contrast of the associated stimulus as:



where *β*=1 for unattended item, *β* >1 for attended items, α is the slope of normalization, *c* is stimulus contrast; *s* is the baseline of normalization.

In the Cortical Microcircuit for Attention model [32], attention was modeled as a change in the driving current to the network neurons. Excitatory (E), feedforward interneurons (FFI) and top-down interneurons (TDI) are differentially modulated by attention: the firing rate of the FFI increases with spatial attention and decreases with feature-based attention, whereas the TDI increase their firing rate with feature-based attention and shift the network synchrony from the beta to the gamma frequency range. The neurons are arranged in columns, each containing neurons of the three types. The synaptic connectivity was designed to ensure inter-stimulus competition and the generation of cortical rhythms. The dynamics of each type of neuron are given by:



where *A* is a scaling factor; *c* is the stimulus contrast; *β* is stimulus selectivity, *0.5 ≤ β < 1*; *I0* is a constant offset current.

The integrated microcircuit model of attentional processing [33] is composed of a reciprocally connected loop of two (sensory and working memory) networks. The channel kinetics are modeled by:





where *s* is the gating variable, *x* is a synaptic variable proportional to the neurotransmitter concentration in the synapse, *ti* are the presynaptic spike times, τ*s* ms is the decay time of NMDA currents, τ*x* controls the rise time of NMDAR channels, and αs controls the saturation properties of NMDAR channels at high presynaptic firing frequencies.

In the predictive coding/biased competition model [34], every processing stage consists of prediction and error detection nodes, with firing rate dynamics defined by:



where *Si* represents the current processing stage, *ySi* and *eSi* the vectors of predictive and error node activations, respectively, ***W*** are matrices of weight values, and *η*, *ς*, and *ϑ* are constant scale values. The external attentional modulation signal is *yAi*, modulating activations through a set of weights ***W****Ai*.