5

Integration and Selection in Multidimensional Dynamic Fields

SEBASTIAN SCHNEEGANS, JONAS LINS, AND JOHN P. SPENCER

This chapter opens Part 2 of the book, which will deal with the integration of lower-level perception-action with higher-level cognition. We begin this chapter with a second look at lower-level visual perception. Recall that Chapters 1–3 introduced several aspects of the neural dynamics of early vision. Here, we'll expand on these concepts, introducing a new type of field—multidimensional dynamic fields. These fields have some concrete advantages. For instance, they enable the fast, flexible *integration* of colors and spatial positions, leading to a neural population response—a peak—that indicates, for example, that the red cup is on the left side of the visual field. There are also some downsides, however: Multidimensional fields are computationally costly; a full multidimensional representation of the visual world would require more neurons than are present in the human brain.

This, in part, motivates a discussion of the flip side to real-time integration—*selection*. Here, we will discuss the advantages of having a system that can selectively "attend" to particular aspects of information captured in a multidimensional field. One advantage is that information can be represented in a simpler, less neurally costly format such as the one-dimensional fields used in Chapters 1–4. Selection also conveys advantages for behavior. For instance, after 10 minutes of cooking dinner in the kitchen, we might not really care that the red cup was on the left side of the visual field. By then, our eyes, head, and body have moved. Instead, we might just want to remember that the red cup was *somewhere* on the counter and then engage in a visual search for that item. This example nicely illustrates the simple forms of "cognition" that we'll tackle here: maintaining a perceptual goal—find the red object—in a self-sustaining or working memory state in a way that can influence early visual processing.

This, of course, is a pretty modest form of cognition. More sophisticated forms of dynamic thinking will emerge across the remaining chapters in Part 2. Critically, however, the insights about integration and selection discussed here set the stage for the discussion of higher-level cognition in subsequent chapters.

NEUROPHYSIOLOGY OF HIGHER-DIMENSIONAL REPRESENTATIONS

In Chapter 2, we introduced the concept of dynamic fields (DFs) using examples that spanned different spaces, including location, direction, and amplitude. Many of these examples involved spaces that span multiple dimensions, such as the visual space that describes two-dimensional positions on the retina. So far, however, we have only treated the mathematics and dynamical properties of one-dimensional DFs. Now we will extend this discussion to multidimensional fields. The space spanned by a higher-dimensional representation may be composed of multiple dimensions of the same type, as in the case of a two-dimensional representation of spatial location, or the representation may combine spaces with qualitatively different features, such as a location dimension and a metric feature like edge orientation or hue value. For both types of representations, numerous examples can be found in neural populations in the brain, reflecting the many multidimensional perceptual and motor spaces that are relevant for the behavior of an organism.

An obvious example for a feature space with multiple qualitatively equal dimensions is physical space. While physical space itself is three-dimensional, we typically perceive spatial aspects of the environment via two-dimensional sensory surfaces, such as the body surface with touch receptors in the skin, or the retinas in the eyes. Correspondingly, we find neural population code representations that span the two-dimensional space of locations on the body surface or two-dimensional visual space. An instance of the latter was already presented in Chapter 3, in the form of neural populations in the superior colliculus, which are involved in the planning of saccadic eye movements. In the neural population forming the superficial layer of the superior colliculus, each neuron has a localized receptive field in retinal visual space, and its activity reflects the presence of salient visual stimuli at the corresponding visual location. The neurons of this population taken together cover the whole visual space with their receptive fields and provide a map of salient stimulus locations. The population activity can therefore most appropriately be modeled by a two-dimensional DF spanning the visual space. This has been employed to model neural activity in the superior colliculus during saccade preparation by Marino, Trappenberg, Dorris, and Munoz (2012).

Many more neural representations of visual space exist throughout the visual-processing hierarchy in the cortex, beginning with the primary visual cortex. Importantly, most of these representations are not purely spatial, but encode additional visual features, such as the orientation of edges in the visual image, movement directions, spatial frequency of visual patterns, and color. One of the earliest probes of visual response properties in primary visual cortex was conducted by Hubel and Wiesel (1959). They found that many neurons, termed *simple cells*, showed a pronounced orientation tuning within their spatial receptive fields. Each such neuron receives (indirect) synaptic input from a small patch of the retina and, due to the specific synaptic connectivity pattern, shows high activity only if the visual image in this region of the retina contains an edge with a particular orientation. The whole population of neurons covers all orientations at all possible retinal locations with their tuning curves (Blasdel, 1992) and forms the basis for the perception of shape and motion. The feature space covered by this neural population is three-dimensional, spanned by two dimensions of visual space and one orientation dimension. Neurons can be effectively positioned within this cube based on their orientation tuning—which orientation they prefer—and their spatial tuning or receptive field—the region in visual space within which they respond to stimulation.

Similar feature maps exist for other visual features, including spatial frequency (Issa, Trepel, & Stryker, 2000), movement direction, color (Livingstone & Hubel, 1984b), and ocular dominance (which is relevant for stereo vision; Blasdel & Salama, 1986). All of these neural populations form representations over a space spanned by two spatial dimensions and at least one feature dimension. To capture the activity of these neural population with a DF model, we can define a field as an activation distribution over this three-dimensional (or higher-dimensional) space. We will use the visual representations over multiple spatial and feature dimensions in the early visual cortex as a central example in this chapter, although we will simplify the spatial aspects of these representations to make it easier to visualize the field dynamics.

Before moving to the dynamics of multidimensional fields, a few comments are in order. First, it is important to stress the difference between the functional dimensionality of a representation as captured in, say, a three-dimensional DF, and the anatomical layout of neurons in cortex. Anatomically, the neurons forming cortical representations are necessarily arranged on the two-dimensional cortical surface (note that subcortical structures may have different anatomical layouts). The feature maps in the early visual cortex, for instance, form a complex, interleaved pattern on the cortical surface, and a great deal of research is aimed at investigating the anatomical layout of such maps (Hübener, Shoham, Grinvald, & Bonhoeffer, 1997; Swindale, Shoham, Grinvald, Bonhoeffer, & Hübener, 2000). We entirely disregard this anatomical layout and only consider the functional dimensionality of a representation. This is motivated by a central goal—to bridge between the functional properties of neural populations in the brain and the dynamics of behavior. We discussed these issues extensively in Chapter 3 with the distribution of population activation (DPA) approach. Recall that the DPA approach was able to capture neural population dynamics within sensory and motor cortex despite the fact that these cortical areas have radically different anatomical layouts (with a topographical organization in visual cortex and a non-topographical anatomical layout in motor cortex). It is important to keep this emphasis on functional topography in mind.

Another key issue that arises with multidimensional fields is how one knows "the functional dimensionality" of a representation. Knowing this amounts, in part, to knowing which dimensions are relevant for a given type of behavior. The neurophysiological experiments just described approached this by varying particular perceptual dimensions (orientation, spatial location) and asking which dimensions affect neural responses. Similarly, behavioral researchers might probe this question by varying particular perceptual dimensions (color, spatial location) and asking which dimensions affect behavioral responses such as reaction times and accuracy. A central characteristic of a multidimensional representation is that neural or behavioral responses systematically vary across the space of possible manipulations. For instance, neurons in visual cortex are responsive to variations in *both* orientation and spatial position and these neural responses capture all possible orientation–space combinations. This is qualitatively different from an assembly of neurons in which one subgroup changes its response depending on stimulus location, and another, separate group responds selectively to certain orientations.

As a final remark, it is worth pointing out that such multidimensional neural representations are costly, much more so than separate low-dimensional representations. This applies both to the computational model and to biological neural systems. Assume, for instance, that we want to form a population code representation for a five-dimensional space. If we want to sample the space along each dimension homogeneously with 100 neurons, the full representation would require a total of 10 billion (100^5) neurons—compared to just 500 neurons to sample each dimension individually. Thus, it is necessary that the neural system is sparing in its use of multidimensional representations. As we will see in the remainder of this chapter, a seemingly high-dimensional representation can be achieved by coupling multiple, lower-dimensional fields together.

THE MATHEMATICS AND DYNAMICS OF HIGHER-DIMENSIONAL FIELDS

Mathematically, the extension of a DF to more than one dimension is straightforward (see Box 5.1 for a formal description). Let's first look at an example with two dimensions of the same type, like a two-dimensional spatial field. In this case, an activation value is assigned to every point in this two-dimensional space, such that we obtain

BOX 5.1 LATERAL INTERACTIONS IN MULTIDIMENSIONAL FIELDS

The general formulation for the differential equation of a DF over a multidimensional space F is as follows:

$$\tau\dot{u}(\mathbf{x}) = -u(\mathbf{x}) + h + s(\mathbf{x}) + \int_F k(\mathbf{x} - \mathbf{x}')\,g(u(\mathbf{x}'))\,d\mathbf{x}'$$

It has the same form as for the one-dimensional field, but the position in the field is now described by a vector, $\mathbf{x} \in F$. If we break up this vector, we can describe the activation of a two-dimensional field as a function of two scalar parameters, x and y. This yields a field equation of the form

$$\tau\dot{u}(x,y) = -u(x,y) + h + s(x,y) + \iint k(x - x', y - y')\,g(u(x',y'))\,dx'dy'$$

A typical lateral interaction kernel (with a Mexican-hat shape) in two dimensions can be described as a difference of two Gaussians, a narrow excitatory component and a wider inhibitory component, with an optional global inhibition term:

$$k(x,y) = c_{\text{exc}} \cdot \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma^2_{x,\text{exc}}} + \frac{y^2}{\sigma^2_{y,\text{exc}}}\right)\right]$$

$$-c_{\text{inh}} \cdot \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma^2_{x,\text{inh}}} + \frac{y^2}{\sigma^2_{y,\text{inh}}}\right)\right] - c_{\text{glob}}$$

Here, c_{exc} is the strength of the lateral excitation, and $\sigma_{x,\text{exc}}$ and $\sigma_{y,\text{exc}}$ are the width parameters along each dimension. Remember that these width parameters may be chosen independently of each other—the interactions may be broad along one dimension, but sharp along the other. The parameters c_{inh}, $\sigma_{x,\text{inh}}$, and $\sigma_{y,\text{inh}}$ analogously describe the inhibitory Gaussian component, and c_{glob} is the strength of global inhibition.

If the field is defined over two qualitatively equal dimensions, as in a representation of two-dimensional physical space, the interaction kernel can be simplified to the form

$$k(d) = c_{\text{exc}} \exp\left(-\frac{d^2}{2\sigma_{\text{exc}}^2}\right) - c_{\text{inh}} \exp\left(-\frac{d^2}{2\sigma_{\text{inh}}^2}\right) - c_{\text{glob}}$$

as a function of the distance in that two-dimensional space, for instance, the Euclidean distance

$$d = \sqrt{(x - x')^2 + (y - y')^2}.$$

Here, the same width parameters σ_{exc} and σ_{inh} are used along both dimensions.

The external input $s(x, y)$ for such a field can in the simplest case be specified using two-dimensional Gaussian patterns. For a single localized stimulus at a location $[p_x, p_y]$, the input can be given as

$$s(x,y) = a_s \cdot \exp\left[-\frac{1}{2}\left(\frac{(x - p_x)^2}{\sigma_{s,x}^2} + \frac{(y - p_y)^2}{\sigma_{s,y}^2}\right)\right]$$

with parameters $\sigma_{s,x}$ and $\sigma_{s,y}$ specifying the width of the stimulus and a_s specifying stimulus strength.

a two-dimensional activation distribution (see Figure 5.1). The field equation now specifies the rate of change of activation within this two-dimensional space. As before, changes in activation are governed, in part, by an interaction function which specifies how local and far-away neighbors influence one another. Critically, the interactions now have to be extended to two dimensions. For two qualitatively similar dimensions (like two spatial dimensions) that share the same metrics, the interaction strength can be defined directly as a function of distance in the two-dimensional space. Analogous to the one-dimensional case, a typical kernel may take the form of a two-dimensional difference of Gaussians (with a "Mexican hat" shape), creating excitation over short distances and inhibition for longer distances. To compute the effect of these interactions, the interaction kernel is now convolved along both dimensions with the field output (the sigmoid function of the activation value at every point in the field). This is then added to the resting level, stimulus inputs, and the current pattern of activation to determine the evolution of activation through time.

Multidimensional fields support the same stable states and instabilities between them as described in Chapters 1–4, provided that the interaction kernels are set up in an analogous fashion (with local excitation and surround inhibition, e.g., as differences of Gaussians). In particular, the interactions promote the formation of localized peaks of activation in the multidimensional space. A detection decision occurs when a peak forms in response to external input, and bistability can be observed when the input strength is subsequently decreased. For sufficiently strong interactions, peaks can become self-sustained. Moreover, strong long-range inhibition can generate a selection behavior in which activated regions compete with each other and a single peak forms in response to multiple localized inputs. And with localized excitatory and inhibitory interactions, multiple peaks can form and, under some conditions, be self-sustained in the absence of input. There are some additional stable states possible in multidimensional fields that do not occur in one-dimensional fields (see, e.g., Owen, Laing, & Coombes 2007), but generally these do not play a role in DF models so we will not discuss them here.

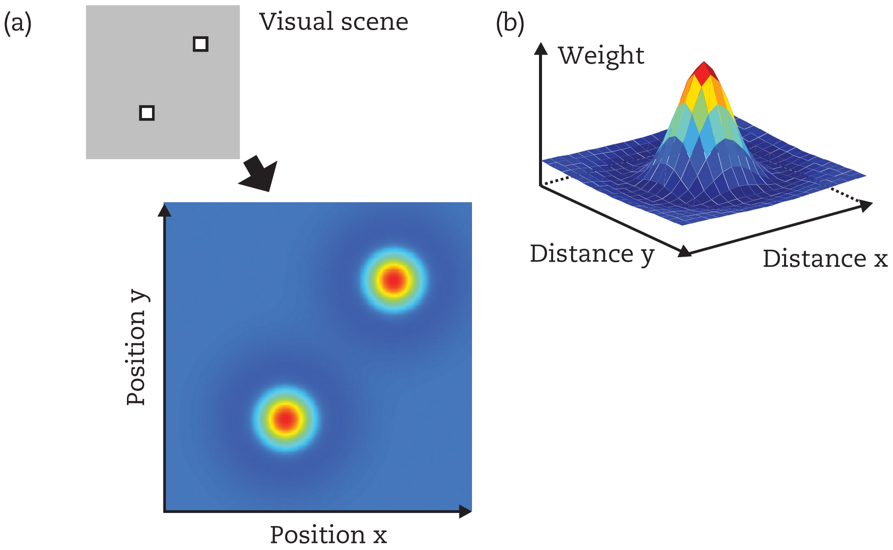

FIGURE 5.1: Two-dimensional field and interaction kernel. *(a)* A dynamic neural field defined over two-dimensional visual space is shown. Activation for each field position is color coded, with dark blue being lowest and red, highest activation. Two stimuli in the visual scene provide localized input to the field, creating activation peaks surrounded by inhibition. *(b)* A difference-of-Gaussians kernel over two-dimensional space. This symmetrical kernel function creates excitation over short distances and surround inhibition over longer distances.

The situation is slightly different for multidimensional fields with qualitatively different dimensions. Let's take as an example a field that spans a spatial dimension along one axis and the angular space of visual edge orientation on the second axis (yielding a simplified model of the orientation representation in early visual cortex). Since a distance in physical space cannot directly be translated into a distance in the space of edge orientations, there is no predefined metric in the combined field that we can use to specify a distance-based interaction kernel. Mathematically, this is a rather trivial issue. We can define any metric for the combined space we like. We can do so implicitly simply by specifying a two-dimensional interaction kernel, for instance, a Gaussian with interaction widths defined for both dimensions (see Box 5.1).

However, we need to consider what the meaning of this metric is. In a nutshell, it specifies which points in the combined space should be considered similar (and thus have excitatory connections between them in the DF model) and which considered dissimilar (and have inhibitory connections between them). In particular, it defines how much the distance within one dimension (such as spatial position) affects similarity compared to the distance along the other dimension (such as orientation). How do we choose a metric that is appropriate for a specific DF model? If we want to simulate a specific neural population and sufficient neural data are available, we can base the metric on neural tuning curves or the pattern of lateral connections within the neural population. Alternatively, we can choose a metric based on functional considerations. It is important to note here that the metric relationship doesn't have to be the same for different DFs defined over the same combined feature spaces. For example, depending on the functional role of a field or the properties of a modeled neural population, it may be appropriate for one field to have broad spatial interactions combined with sharp interactions for the orientation dimension, and vice versa for another field. In practice, the critical point is that the specific form of the interactions across each dimension leads to behavioral and neural signatures that are, in principle, observable. Thus, even in fields with qualitatively different dimensions where the metrics can have an arbitrary feel, this is ultimately not the case once the mapping to the brain and behavior are considered.

Fields that span more than two dimensions can be formed in an analogous fashion: The field is defined as an activation distribution over the higher-dimensional space, an interaction kernel in the same space is defined, and interactions are computed as the multidimensional convolution of the kernel with the gated field output. The number of possible dimensions is only limited by the increasing computational load and by considerations of the neural plausibility of population representations over very high-dimensional spaces.

REAL-TIME INTEGRATION AND SELECTION IN DYNAMIC FIELDS

In this section, we explore the functional role of multidimensional DFs using feature maps in early visual processing as examples. For now, we only consider a single visual feature, namely color. More precisely, we represent color through an activation distribution over the space of hue values (see, e.g., Wachtler, Sejnowski, & Albright, 2003, for the neural representation of color in early visual cortex). A full DF model of the feature maps in early visual cortex would then span two spatial dimensions and this hue dimension. While we do employ such three-dimensional DFs in both behavioral models and robotic architectures (see Chapters 7 and 9), here we simplify the model by considering only one spatial dimension. In particular, we assume that all stimuli are presented along a horizontal line and we can, therefore, ignore the vertical position. The resulting two-dimensional field (over one horizontal spatial dimension and one color dimension) retains all the functional properties of the full model but allows an easier description and illustration of activation patterns.

As a convention for the depiction of this field, we display the spatial dimension on the *x*-axis with the color (hue) dimension on the *y*-axis (Figure 5.2). This field receives external input—visual input on the retina combined with early visual processing—that is localized along both axes. The presentation of a green spot of light in the left half of the visual field in Figure 5.2, for instance, creates an input localized in the left part of the field with a vertical position that reflects the hue value "green." If this input is strong enough, it creates a localized activation peak that is stabilized by lateral interactions. Given that this field reflects processing in early visual cortex, we do not use strong neural interactions with global competition. Consequently, if a second stimulus is presented—a blue spot of light on the right side of the visual field (see Figure 5.2)—it produces another localized activation peak that is separated from the first one both along the spatial and the color dimension. When a stimulus is turned off, the corresponding activation peak goes through the forgetting instability and returns to the resting state.

We now add two one-dimensional fields to this system. The first is a purely spatial representation,

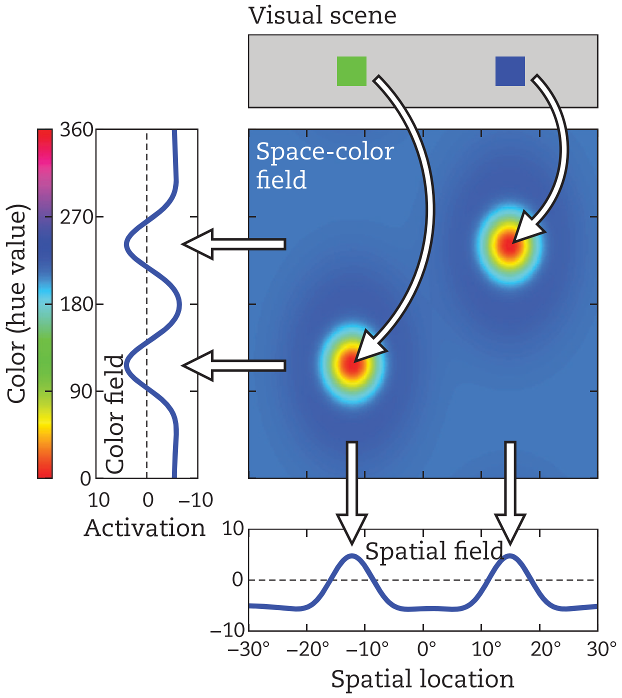

FIGURE 5.2: Read-out from space-color field to separate one-dimensional fields. The two-dimensional field is defined over the combined space of color (hue value, vertical axis) and horizontal spatial location (horizontal axis). It receives localized visual inputs that reflect the combination of color and location for each stimulus. These stimulus features can be read out into separate one-dimensional fields for color and spatial location, shown aligned with the corresponding axis of the two-dimensional field. The read-out is performed by integrating the field output over the disregarded dimension.

spanning only the spatial dimension from the two-dimensional field, and the other is a color representations that spans the dimension of hue values (Figure 5.2). In the illustrations, the feature axes of these two fields are aligned with the corresponding axes of the two-dimensional field such that the effects of projections between the two fields can be seen more clearly.

We will assume here that only the two-dimensional field receives direct visual input. This reflects the situation in the nervous system. The earliest visual representations in the cortex are specific to both surface features and their locations. As described earlier, they are composed of neurons that act as localized detectors for simple features and show high activity if, for example, a certain orientation or a certain color is present at a specific location in the visual image. Representations like the one-dimensional color field instead reflect the presence of a color in the visual image independent of its location and can be seen as a higher level of abstraction. Accordingly, the one-dimensional fields in this architecture receive their input from the two-dimensional field. As shown in Figure 5.2, these one-dimensional fields enable us to "read out" the positions and colors of all stimuli in the scene. The (horizontal) positions of all stimuli are given by the peak locations in the one-dimensional spatial field, while the colors of the stimuli are given by the peak locations in the one-dimensional color field.

The basic operation that enables this form of "read-out" is to integrate (i.e., sum up) the output over the disregarded dimension (see Box 5.2 for a formal description of this operation). That is, to extract the spatial locations of all stimuli, we integrate the output of the two-dimensional field over the color dimension. The result is close to zero for all spatial locations that do not contain any peaks, and significantly different from zero only around the peak positions. This integrated output provides the input to the one-dimensional spatial field. It is, however, not fed in directly. Rather, this input is first smoothed with another Gaussian kernel, reflecting our assumption that projections are not point-to-point but are smoothed by the synaptic spread found in neural populations. The input strength is also scaled

BOX 5.2 DYNAMIC FIELD ARCHITECTURE FOR SPACE-FEATURE INTEGRATION

The basic architecture for the integration of space and surface features described in this chapter consists of three fields: the two-dimensional visual sensory field, $u_{\text{vis}}(x,y)$, defined over one spatial and one feature dimension; the one-dimensional spatial field, $u_{\text{spt}}(x)$; and the one-dimensional color field, $u_{\text{col}}(y)$. For the projection from the visual sensory field to either the spatial or the color field (the "read-out"), we need to integrate the field output over the disregarded dimension. The integrated output is then convolved with a one-dimensional Gaussian interaction kernel. For the spatial field, this yields the field equation

$$\tau\dot{u}_{\text{spt}}(x) = -u_{\text{spt}}(x) + h_{\text{spt}} + s_{\text{spt}}(x) + \int k_{\text{spt,spt}}(x-x')\, g\left(u_{\text{spt}}(x')\right)dx' + \int\left[k_{\text{spt,vis}}(x-x')\int g\left(u_{\text{vis}}(x',y)\right)dy\right]dx'$$

The last term in this equation describes the input from the visual sensory field. The inner integral computes the visual sensory field's output for one spatial position, integrated over all feature values. The outer integral then describes the convolution with the interaction kernel $k_{\text{spt,vis}}$ (the index here specifies that this kernel mediates the projection from the visual field to the spatial field). Analogously, the field equation for the color field reads

$$\tau\dot{u}_{\text{col}}(y) = -u_{\text{col}}(y) + h_{\text{col}} + s_{\text{col}}(y) + \int k_{\text{col,col}}(y-y')\, g\left(u_{\text{col}}(y')\right)dy' + \int\left[k_{\text{col,vis}}(y-y')\int g\left(u_{\text{vis}}(x,y')\right)dx\right]dy'$$

Here, the output of the visual sensory field is integrated over the spatial dimension (inner integral in the last term of the equation), and the result is convolved with the Gaussian interaction kernel $k_{\text{col,vis}}$ along the color dimension.

The visual sensory field in turn receives ridge inputs from the spatial field and the color field, which are also smoothed with Gaussian interaction kernel. These ridges can be expressed in the field equation in a relatively simple fashion, as inputs that depend only on either the x-position or the y-position within the two-dimensional field:

$$\tau \dot{u}_{\text{vis}}(x,y) = -u_{\text{vis}}(x,y) + h_{\text{vis}} + s_{\text{vis}}(x,y) + \iint k_{\text{vis,vis}}(x-x', y-y')\, g\big(u_{\text{vis}}(x',y')\big)\, dx'dy' \\ + \int k_{\text{vis,spt}}(x-x')\, g\big(u_{\text{spt}}(x')\big)\, dx' + \int k_{\text{vis,col}}(y-y')\, g\big(u_{\text{col}}(y')\big)\, dy'$$

The second to last term specifies the input from the spatial field, the last term the input from the color field. The double integral term describes the lateral interactions in the two-dimensional field, as detailed in Box 5.1.

in such a way that the state of the spatial field is not completely dictated by the input. This allows modulations of, for instance, the resting level in the spatial field to impact whether this field builds one peak or multiple peaks from the input it receives. The color information can be extracted from the two-dimensional field in an analogous fashion by integrating the field output over the spatial dimension. This summed output is likewise smoothed and scaled and fed as input into the one-dimensional color field.

Let's take a look at the resulting representation in the complete system (Figure 5.2). If the visual stimulus consists of the green point of light on the left and the blue point of light on the right, we have two localized peaks in the two-dimensional field. There are likewise two activation peaks in each of the one-dimensional fields. The spatial field has one peak on the left and one on the right, reflecting the stimulus positions; the color field has one peak at the hue value for "green" and one peak at the hue value for "blue." At this level of description, then, the two one-dimensional fields reflect the same spatial and color values represented in the two-dimensional field, and they do so in a much more compact form. Assume that in the DF model we sampled each dimension with 100 discrete nodes. In this case, the two one-dimensional fields can be modeled with a total of 200 nodes, while the single two-dimensional field requires $100 \times 100 = 10{,}000$ nodes to sample with the same resolution.

So what is the advantage of having the two-dimensional field, besides the fact that it may more accurately reflect the actual neural representation in the visual cortex? One key piece of information that is lost in the one-dimensional fields is the *integration of colors and space*—the particular combinations or "bindings" of spatial and color values. Concretely, with only the information from the one-dimensional fields, one can tell that there is a blue and a green item, and that one item is on the left and one on the right. One cannot tell, however, whether the green item is on the left or on the right. To show this, imagine that the two items in Figure 5.2 switched their locations. Now the peaks in the two-dimensional field will reflect the change as both activation peaks move to previously unoccupied locations. The activation patterns in the one-dimensional fields, however, remain unchanged—there is still a blue and a green item in the scene, and the same two spatial locations are occupied.

Knowing the locations of specific features and, likewise, the conjunction between features is critical for many aspects of goal-directed action. If you want to reach for an apple, for instance, it is not sufficient to know that there is a red item in the visual scene along with a blue and yellow one, and that one of them is to the right, one in the middle, and one to the left. You have to know which of the items actually is the red one to make the correct movement. This requires the real-time integration of color and spatial position captured by the two-dimensional field.

On the other hand, behaviors do not always depend on fully "bound" information; sometimes only a subset of the possible perceptual dimensions is relevant. Once the right target is selected, for example, the neural system that drives the actual planning and control of the reach movement does not need to know that the item is red. It is, therefore, desirable to have a purely spatial representation to stabilize this aspect of behavior, because reaching behaviors are invariant with respect to color.

Similarly, when you think about the apple and conjure up its image, you do not have to link this image to a specific location. Thus, invariance to object location is often a desirable feature of behaviors such as object recognition.

These examples highlight the yin and yang of integration and selection in neural systems. Integration is critical when specific feature combinations for multiple items are simultaneously present or must be simultaneously maintained in, for instance, working memory (a case we will return to in Chapters 8 and 9). This is certainly the case in early visual processing where multiple items are likely to be simultaneously present in the visual field. Knowing which features correspond to which objects is absolutely critical, as is doing this quickly and flexibly. Indeed, it is worth emphasizing that the two-dimensional field can quickly build a peak for *any* combination of (horizontal) spatial position and color. The flip side to integration is selection. Selection is critical to avoid computational and neural overload and to stabilize particular "local" decisions. This was the case in the reaching example when decisions about a spatial location—and not about color—were important for planning a reach.

Given that both integration and selection are critical to visual cognition, how do we efficiently use both types of neural representation in concert, avoiding duplication of effort on the one hand and errors on the other? For instance, how do we avoid scenarios like the one just described, where we don't know whether the green item is on the left or the right? A simple way to avoid this problem is to force all fields to have a single peak. In this case, the single activation peak in each one-dimensional field informs us of the color and the position of the stimulus, and the activation pattern in the two-dimensional field does not add any further information. We can also add another spatial dimension and more feature dimensions (e.g., orientation, spatial frequency) and still get the full information from the separate one-dimensional fields. That is, we can know that the red, shiny, apple-shaped object is near the upper left corner of the desk.

Of course, this scenario does not take advantage of the multiple, integrated representations possible in the two-dimensional field, nor would it work effectively in a real, cluttered visual world where multiple items are simultaneously visible. But the considerations here lead to an insight: By selecting perceptual items one at a time in the one-dimensional fields, we can make use of these lower-dimensional representations in many situations, while still avoiding "binding" errors. This can be viewed as an implementation of selective attention (Desimone & Duncan, 1995). If the selectivity is primarily driven by attention to a single feature value, we would refer to this as feature-based attention. Reversely, selectivity based on attention to a spatial value would reflect spatial attention.

To implement this kind of coupling, we need connections that project back from the lower-dimensional fields to the higher-dimensional ones (the opposite of the "read-out" connections described earlier). For instance, we want the one-dimensional color field in our example system to project back and influence activation in the two-dimensional field over color and space. Since the color field does not contain any information about stimulus position, its input to the combined field should be homogeneous along this dimension. The one-dimensional field does, however, provide specific color information, so it should activate the same region along the color axis in the two-dimensional field that is active in the color field. The resulting input pattern takes the shape of a horizontal ridge of activation running through the two-dimensional field (Figure 5.3a; see Box 5.2 for a formal description of this operation). Typically, this ridge input should not be so strong that it can induce a peak on its own (since the peak would form at a random location along the ridge). Rather, the strength should be sufficient to induce a peak in concert with other inputs. The projection from the one-dimensional spatial field to the two-dimensional field can be defined analogously: this input drives activation locally along the spatial dimension, but is unspecific for color and, therefore, generates a vertical activation ridge in the two-dimensional field (Figure 5.3b).

With bidirectional connections between each one-dimensional field and the combined two-dimensional field in place, we can use the DF architecture to solve some simple visual tasks. For the scenarios discussed later in this chapter, only the two-dimensional field receives direct visual input. The one-dimensional fields can receive external inputs (e.g., a global boost instructing the system to "attend" to color or spatial position), but these will be from unspecified "top-down" sources at present, that is, we'll implement these "top-down" influences by hand (later in the chapter, we'll provide a detailed account of such processes). Further, we set the interactions in the one-dimensional fields such that these fields are selective (the single-peak

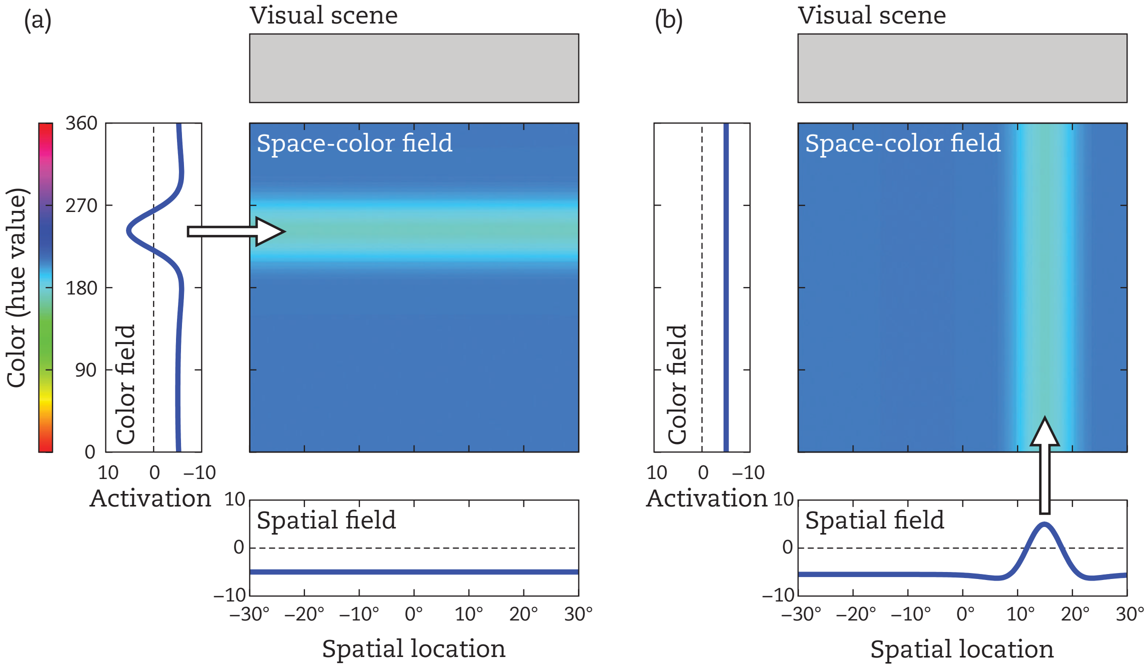

FIGURE 5.3: Ridge input from one-dimensional fields to a two-dimensional field. (*a*) The projection from the one-dimensional color field provides specific (localized) input along the color dimension but is homogeneous along the spatial dimension. This produces a horizontal ridge of subthreshold activation in the two-dimensional field over the combined space of color and location. (*b*) The input from the one-dimensional spatial field conversely is unspecific for color and consequently produces a vertical activation ridge.

regime). We also use a relatively weak input from the two-dimensional field into the one-dimensional fields. Thus, if there are multiple visual stimuli in the scene, each of them will create an activation peak in the two-dimensional field, but the one-dimensional fields will show only subthreshold "bumps" of activation in response to these inputs.

We begin with a simple demonstration. Assume our system looks somewhere in the world and comes upon a set of visual items. The inputs to the two-dimensional field will create peaks of activation that reflect the colors and spatial positions of the items (Figure 5.4a). In addition, the one-dimensional fields will have hills of subthreshold activation at the associated positions and color values. Now, let's apply some top-down control by boosting the spatial attention field homogeneously (Figure 5.4b). This will lead to the selection of one of the positions currently stimulated by input from the two-dimensional field. The spatial selection will be random if all stimuli are of equal strength; the selection will be biased if one item is more salient (or more intense) than the others. Once the position is selected in the spatial field (i.e., a peak is built), this field will project input back into the two-dimensional field at the associated spatial position. This will boost one of the peaks in the two-dimensional field far enough to drive up activation in the color field at the associated hue value. We can then either set the "baseline" resting level of the color field to build a peak under these conditions or we can extract the color at the selected location by boosting the color field (i.e., by "attending" to the associated color).

We can also implement a more "local" type of top-down influence. Say we want to look for a blue item. In this case, we can implement a top-down influence by providing a direct input to the color field that induces a peak at the "blue" location along the color axis (Figure 5.5). The projection from the color field to the two-dimensional field then induces a horizontal ridge input for that color. Assuming that one of the stimuli in the visual scene is indeed blue, there is already an activation peak in the two-dimensional field that lies right on this input ridge. This peak will be strengthened by the ridge input and become larger than the other peaks in the field. Since this peak also projects back to the color field, the two peaks will reciprocally excite each other and grow further, but the lateral inhibition should limit the spread of activation. In addition, the amplified peak in the two-dimensional field will also project more strongly to the spatial field. Consequently, the position of the blue item

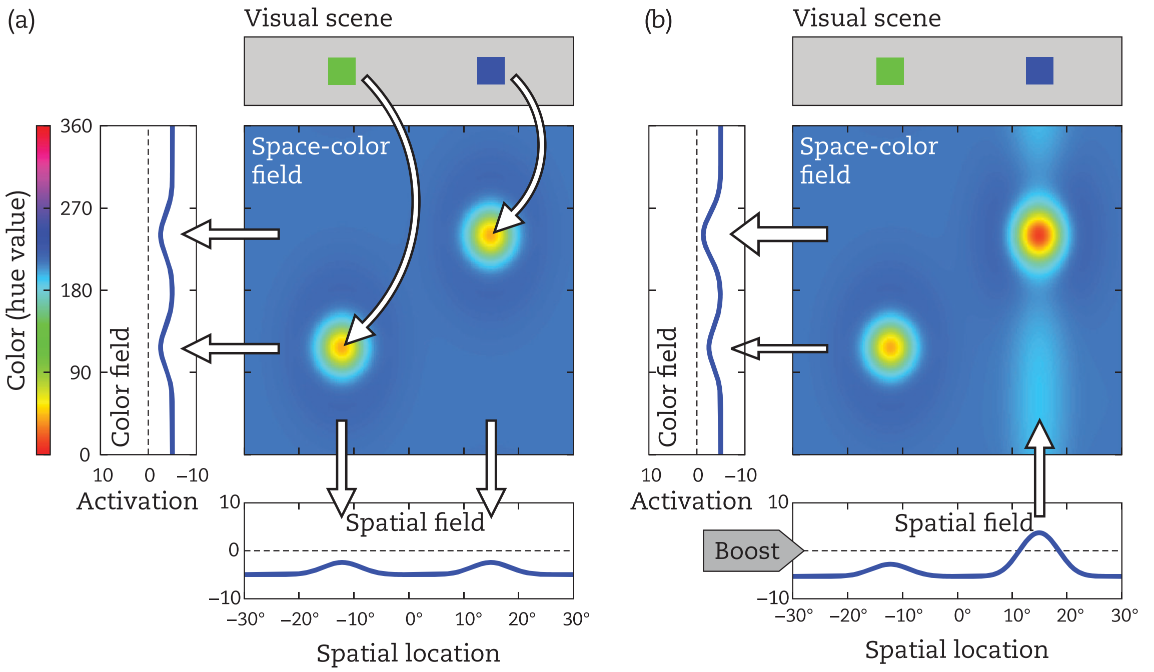

FIGURE 5.4: Spatial selection of a visual stimulus. (*a*) Visual stimuli induce peaks in the two-dimensional space-color field, which in turn project to the one-dimensional fields for space and color. This is analogous to what is shown in Figure 5.2, but here the strength of the read-out projections is reduced such that the activation in the one-dimensional fields remains subthreshold. (*b*) The spatial field receives a global excitatory "boost" input. This lifts the activation for the two stimulus locations in this field beyond the output threshold and initiates a selection process due to the competitive lateral interactions. Here, the location on the right has prevailed and an activation peak has formed, which projects a ridge input back into the space-color field. As a consequence, the input-driven activation peak for the right stimulus is strengthened in the space color field and projects more strongly to the color field. This difference in projection strength can then be used to determine the color of the spatially selected item in a subsequent step.

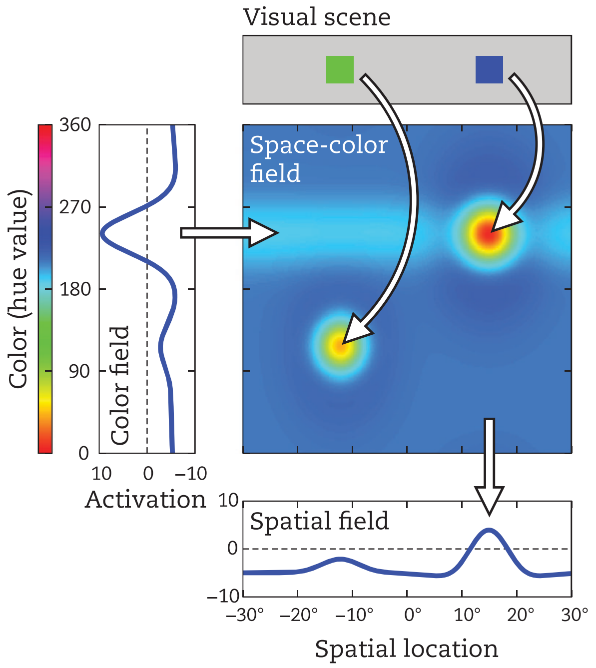

FIGURE 5.5: Selection of a spatial location for a given color. Here, localized visual input to the two-dimensional field is combined with a ridge input from the color field. A desired color is set in the one-dimensional color field through an external input, creating an activation peak at the position for "blue." The color field then projects a horizontal ridge into the combined space-color field. This two-dimensional field also has formed two localized peaks, reflecting the present visual stimuli. The input ridge from the color field crosses one of these peaks and amplifies it. When the spatial field is now boosted homogeneously, the spatial location of this item is selected.

receives stronger input than the positions of all other stimuli in the visual scene. When we then boost the spatial field to select a location, neural interactions in the spatial field will lead to the formation of a single peak at the location of the strongest input. Therefore, we will most likely create a peak on the right at the location of the blue item (see Figure 5.5).

The same mechanism can also be applied in the opposite direction. Say that instead of looking for a blue object we want to figure out what object is over to the left. We start with the same setting as before, but now we first induce a peak at the predefined location in the spatial field (i.e., "left"). This peak induces a vertical ridge input in the two-dimensional field, which amplifies any activation peaks that it overlaps. When the color field is then boosted, it forms a peak at the location of the strongest input, yielding the color of the spatially selected item.

In these examples, we have assumed that there is always a single peak located right on the ridge input. This doesn't always have to be the case. For instance, one may be looking for a red object, but there is none present. There may be similar items however, reflected by peaks in the two-dimensional field that are close to the position for "red" along the color dimension. Since the input ridge is not perfectly localized on a single discrete color, such peaks will still overlap with the ridge input and be strengthened by it to a certain degree. It may also happen that multiple items in a scene match the desired color (either perfectly or partially), so that there are multiple peaks supported by the ridge input.

The response behavior of the system in such scenarios depends on the detailed settings of the connections between the fields and the boost that is applied to generate the response. On the one end of the spectrum, we can imagine a system with strong competitive interactions and a strong external boost in the spatial attention field. Such a system will always produce a single location as a response, typically the location of the best matching item, or a random location if no stimulus matches the desired color at all. On the other end of the spectrum, with a relatively weak boost and no competition in the spatial attention field, a peak may form only if there is a good enough match in the visual sensory field that will provide a strong input to the spatial attention field (the response peak will than form through a detection instability rather than a selection instability). In this latter case, multiple peaks may form as a response if there is more than one matching item in the scene. These different behaviors may (to some degree) also emerge in a single system depending on different modulatory inputs.

The scenarios used thus far are closely tied to the details of early visual processing. In the sections that follow, we expand upon these scenarios by embedding them within particular research paradigms. This highlights both how research paradigms can be useful in expanding a simple neural architecture, and how specific neural models can be constrained by experimental findings and inform the search for new phenomena and the generation of novel predictions.

INTEGRATION AND SELECTION IN AN AUTONOMOUS VISUAL EXPLORATORY SYSTEM

The examples presented thus far all present snapshots in time of, for instance, how the visual system attends to a location to select the feature present there. In this section, we generalize these processes by embedding them within a broader neural system that can actively maintain a perceptual goal (e.g., a working memory for a color), perceive features in space, attend to those features and locations, *and move its "eyes"* (e.g., its virtual camera). This enables the system to move beyond one interval of time—beyond one selection decision—toward cycles of perceiving, attending, remembering, and looking. In this sense, one goal of this section is to close the loop on perception and action to create a more embodied and autonomous view of visual processing and exploration.

To provide a little background for this model, let's quickly review some properties of the human visual system. First, the distribution of light-sensitive neurons over the area of the retina in the human eye is very uneven. The cone cells in particular, which provide color information and high-acuity spatial information, are highly concentrated in the fovea at the center of the retina. A correspondingly large portion of the visual cortex is dedicated to process the sensory information from this foveal region. In order to perceive the fine details of a visual stimulus, it is necessary to fixate it, that is, to bring its image into the fovea. This is achieved by saccades, rapid eye movements of which humans can make several per second. (In Chapter 7, we will address the problem of how visual spatial representations can be kept stable despite the constant shifts of the retinal image.)

Saccadic eye movements are strongly coupled to spatial attention. If a certain location in the visual scene captures our attention (e.g., by a movement in that region), we will likely make a saccade to bring it into the fovea and perceive it in greater detail. This coupling between attention and saccade planning is not absolute, however. It is possible to intentionally suppress saccades and to attend to a location without fixating it.

The specific model we describe here—which we will refer to as the *biased competition* model—was motivated by decades of work examining the interactions among visual working memory (VWM), visual attention, early perceptual processing, and saccade planning and execution. These interactions have been examined from diverse vantage points. For instance, in models of visual search, VWM is proposed to be the substrate for a search template, serving to guide attention and gaze to objects that match task-relevant features in the array (Hamker, 2005). Similarly, models of attentional selection propose that competitive interactions among stimuli are "biased" by signals from VWM (Deco & Lee, 2004). And models of transaccadic memory and integration depend on VWM to span saccades and integrate features across perceptual disruption (Fazl, Grossberg, & Mingolla, 2009). Across these literatures, data suggest that visual attention controls the information initially encoded into VWM (Irwin, 1992; Irwin & Gordon, 1998; Schmidt, Vogel, Woodman, & Luck, 2002), and attention precedes the eyes to the next saccade target (Hoffman & Subramaniam, 1995; Kowler, Anderson, Dosher, & Blaser, 1995). The content of VWM, in turn, modulates the selection of objects as the targets of attention and gaze.

The biased competition model implements particular hypotheses regarding how VWM interacts with attention and early visual processing and influences the most rapid and elementary forms of saccadic orienting. The starting point for this model is the DF model used previously with a two-dimensional visual sensory field, a one-dimensional spatial attention field, and a one-dimensional feature attention field. We expand this model in two directions here, enriching the spatial pathway to capture the planning and execution of saccadic eye movements and enriching the feature pathway to capture aspects of visual working memory.

The model architecture is generally inspired by the two-streams hypothesis (Mishkin & Ungerleider, 1982). This hypothesis posits that visual information is processed in two at least partly independent streams, the "where" and the "what" stream. The "where" (or dorsal) stream deals primarily with the spatial aspects of visual information that form the basis for movement planning (including eye movements). The "what" (or ventral) stream is focused on surface features used for object recognition, and neurons in the higher areas along this path show very large spatial receptive fields, corresponding to low spatial selectivity (Desimone & Gross, 1979). We will return to this hypothesis in Chapter 8 and discuss it in greater detail there. Architectures similar to the one proposed here have been used to model visual search, for instance by Hamker (2004, 2005). In these papers, a subset of cortical areas involved in visual search was modeled, namely the frontal eye field (as a purely spatial representation, but limited to six discrete locations), the inferotemporal cortex (as a purely feature representation), and area V4 (as a combined spatial and feature representation that receives visual input).

In accordance with the principles described in previous chapters, we refrain from claiming such a strict one-to-one relationship between individual neural fields and cortical areas for our model. Instead, we propose that the fields reflect certain aspects of distributed neural representations. In this view, the one-dimensional fields of the feature pathway in the model reflect the pure feature aspects of neural representations in the "what" stream, including V4, the inferotemporal cortex, and contributions from the prefrontal cortex for working memory. The one-dimensional spatial fields model aspects of several cortical and subcortical areas involved in spatial attention and saccade control, including posterior parietal cortex, frontal eye field, and superior colliculus. The two-dimensional field reflects early visual representations with relatively small spatial receptive fields and sensitivity for simple surface features, as found in areas V1 to V4.

The full model architecture is shown in Figure 5.6. The spatial pathway consists of two fields spanning the space of horizontal retinal position and mirrors previous neural field models of saccadic movement planning (Kopecz & Schöner, 1995; Trappenberg, Dorris, Munoz, & Klein, 2001; Wilimzig, Schneider, & Schöner, 2006). The spatial attention field is as before. This field receives direct input from the two-dimensional visual sensory field (integrated over the color dimension) and projects back to it. Lateral interactions in this

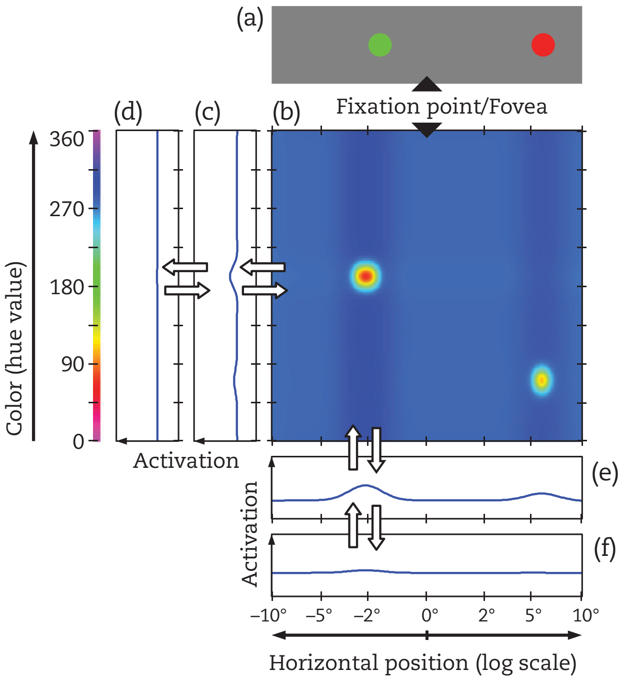

FIGURE 5.6: Architecture of the DF model of working memory influences on saccade planning. (*a*) Visual scene, with colored stimuli. (*b*) Visual sensory field. (*c*) Color attention field. (*d*) Color memory field. (*e*) Spatial attention field. (*f*) Saccade motor field. See text for further explanation.

field implement a soft winner-takes-all mechanism: If multiple distant regions are activated, they compete with each other by means of localized self-excitation and global inhibition, typically leading to the selection of a single area that receives salient visual input. The feedback to the visual sensory field increases that field's activation locally along the spatial dimension but is homogeneous along the color dimension. It thus implements a form of spatial attention. To capture the characteristics of the human visual system, the spatial dimension in the two-dimensional field and in the two fields of the spatial pathway is scaled logarithmically. Thus, the foveal region at the center of the field has an increased spatial resolution and a higher responsiveness, especially to small stimuli.

The second field in the spatial pathway is new—the saccade motor field. This field receives input from the spatial attention field and has strong local excitatory and global inhibitory interactions. If the spatial attention input drives the field activation over the output threshold, the saccade field quickly forms a strong activation peak at the stimulated location and suppresses the activation in all other parts of the field. The peak in this field is the movement plan for the saccadic eye movement.

To actually execute the saccade, we add a saccade motor system. In this system, the amplitude of the saccade is scaled by the position of the peak in the field—the farther the peak is from fixation, the larger the saccade amplitude. (Note that amplitude is the critical dimension in our examples because the task only involves eye movements along a horizontal dimension.) The formation of an activation peak in the saccade motor field initiates a saccadic eye movement. The peak directly produces a motor signal as long as it persists, scaled according to the distance of the peak from the foveal region in the field (with a stronger signal as this distance increases). At the same time, the field output also provides input to a saccade reset node (simulating a pool of homogeneously connected neurons). This input is not scaled with the peak position, so the activation of the saccade reset node rises with a largely fixed time course and reaches the output threshold after a certain duration. In effect, this node acts as a simple neural integrator that measures the total output of the field. Once the node reaches threshold, it strongly inhibits the whole saccade motor field, extinguishes the peak there, and thereby terminates the saccade. While the saccade is in progress, the external input to the visual sensory field is suppressed. When the visual input is reinitialized after the completion of the saccade, it is shifted to reflect the new gaze direction.

The two fields of the feature pathway are defined over the dimension of color (hue value). As in the previous examples, the color attention field receives input from the visual sensory field along the color dimension (integrating over all spatial positions) and projects modulatory input back to it, implementing a feature attention mechanism. This field is bidirectionally coupled to a new field—the color memory field. The coupling between these fields enables memorized content to influence attentional processing, while keeping attentional processing open to visual stimulation from the visual sensory field. As with the color attention field, lateral interactions in the color memory field are implemented with local self-excitation and surround inhibition. These interactions are stronger, however, in the color memory field, such that peaks can be actively sustained even if the input that induced the peak is no longer present. This implements a neurally grounded form of working memory. Note that this self-sustaining activation in the memory field is dependent on a global modulatory input from a memory control node. This node implements a form of "top-down" control, modulating whether peaks in the working memory layer should be actively maintained or simply forgotten.

VISUAL WORKING MEMORY AND SACCADE ORIENTING IN THE REMOTE DISTRACTOR PARADIGM

To understand how the biased competition model integrates VWM, attention, early perceptual processing, and saccade orienting, it is useful to consider the model in a specific behavioral paradigm. This will help ground discussion of the model; it also serves the function of showing how a DF model can be used to explain—in quantitative detail—behavioral performance from well-controlled experimental paradigms.

We focus here on one case study—the remote distractor paradigm from Hollingworth and colleagues (Hollingworth, Matsukura, & Luck, 2013). The design of this paradigm is shown in Figure 5.7. Each trial began with fixation of a central cross and the presentation of a color memory square for 300 ms. The memory stimulus was followed by a blank delay of 700 ms. Next, the saccade target and a distractor were presented. Participants were instructed to execute a saccade to the target as rapidly as possible. They were also instructed that the distractor was irrelevant to the task, and they should avoid fixating it (to emphasize this, the distractor was only present on some trials). The target always appeared much further from central fixation than the distractor and was larger than the distractor, removing any significant ambiguity in the task of orienting to the target. After participants fixated the target, the target display was replaced with the memory test stimuli, which consisted of two color squares. Participants indicated which of the two test squares matched the color of the memory square presented at the beginning of the trial.

Schneegans, Spencer, Schöner, Hwang, and Hollingworth (2014) used this paradigm to

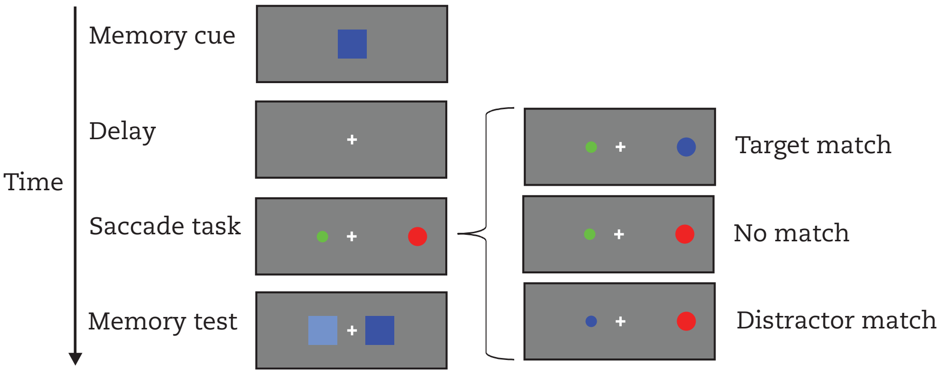

FIGURE 5.7: Remote distractor paradigm testing the influence of color working memory on saccade planning. The paradigm combines two tasks, a working memory task and a saccade task. The working memory task consists of presentation of a color stimulus at the beginning of each trial, and a memory test at the end of the trial in which the subject has to select the memorized color from two alternatives. During the memory period, a saccade task is executed, in which the subject has to make a timed saccade to a target stimulus while suppressing saccades to a simultaneously appearing distractor. The target is distinguished from the distractor by its larger size and its position, always appearing farther from the fixation point than the distractor. The colors of the target and the distractor stimulus are varied to test influences of color memory on saccade planning: Either the target item, the distractor item, or neither of them matches the currently memorized color.

examine how VWM and feature-based attention interact during the planning of orienting saccades in this very simple task. A central manipulation from this study is shown in the right panel of Figure 5.7: on some trials, the target matched the color held in memory; on other trials, the distractor matched the color held in memory; and on still other trials, none of the items matched the color held in memory. Behavioral results are shown in Figure 5.8. Even though the distractor was irrelevant to the task and it was always presented in the opposite direction relative to the target, participants still looked to the distractor on some of the "no-match" trials. Most critically, orienting to the distractor was systematically modulated by the contents of VWM. When the distractor matched the color in VWM, participants looked toward the distractor on nearly half of the trials. By contrast, when the target matched the color in VWM, participants almost never looked at the distractor.

The biased competition model allows us to describe the real-time neural processes that underlie performance in this task, including the formation and maintenance of a color memory, the perceptual processing in response to the target and distractor stimuli, and the specification of a saccadic eye movement. Here, we give a detailed description of this model, unpacking the processes that underlie performance in the remote distractor paradigm. We then quantitatively compare the model's performance relative to data from Schneegans et al.

Each simulated trial begins with the presentation of the memory stimulus at the initial fixation point (Figure 5.9a). This stimulus produces a strong activation peak in the visual sensory field, which is located in the foveal region (at the field center) along the spatial axis, while the position along the color axis reflects the stimulus color (a red hue). The projections along the feature pathway induce a peak for the stimulus color in the color attention field, which in turn provides input to the color memory field. During this phase of the trial, the memory control node is strongly activated to reflect the task instruction to memorize the color of this stimulus (the cognitive processes leading to this activation are not explicitly modeled). This node projects global excitatory input to the color memory field, thereby lifting the memory field's overall activation level. This allows the field to form an activation peak at the position of the localized input from the color attention field and memorize the color of the red item.

The stimulus also elicits activation in the spatial pathway, which projects activation back onto the visual sensory field and further strengthens the activation in the foveal region. It does not, however, elicit a saccadic eye movement because the central section of the spatial attention field does not project to the saccade motor field (since an eye movement is not needed when an item is currently in the foveal region). Instead, the peak in the spatial attention field acts to stabilize fixation by suppressing activation in peripheral regions through global inhibitory interactions.

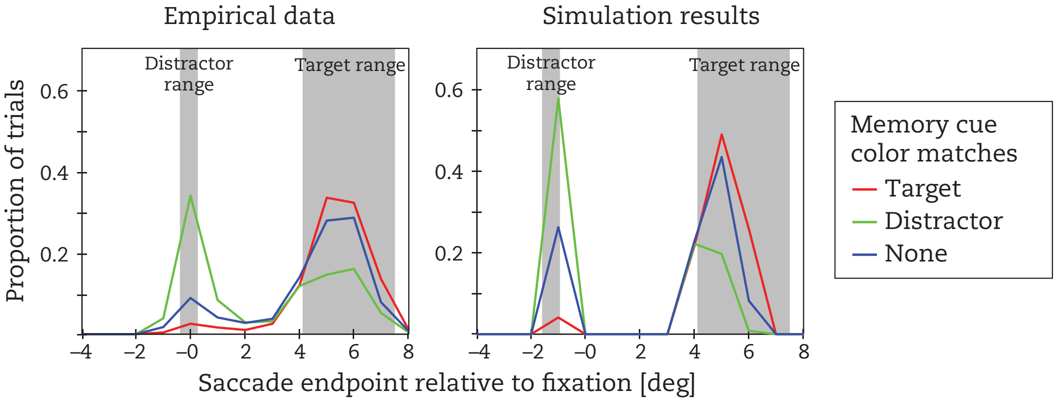

FIGURE 5.8: Empirical and simulation results for the remote distractor paradigm. The histograms show the distribution of saccade endpoints in the saccade task for three different conditions: if the memorized color matches the color of the target, the distractor, or neither of them. In all three conditions, there is a bimodal distribution with a part of the saccades landing near the distractor and a part in the range of the target. The proportion of saccades to the distractor changes significantly between the three conditions: If the target matches the memorized color, there are nearly no saccades to the distractor; if the distractor matches, nearly half of the saccades land on or near the distractor location. If neither matches, there is an intermediate number of saccades to the distractor. The DF model reproduces these results.

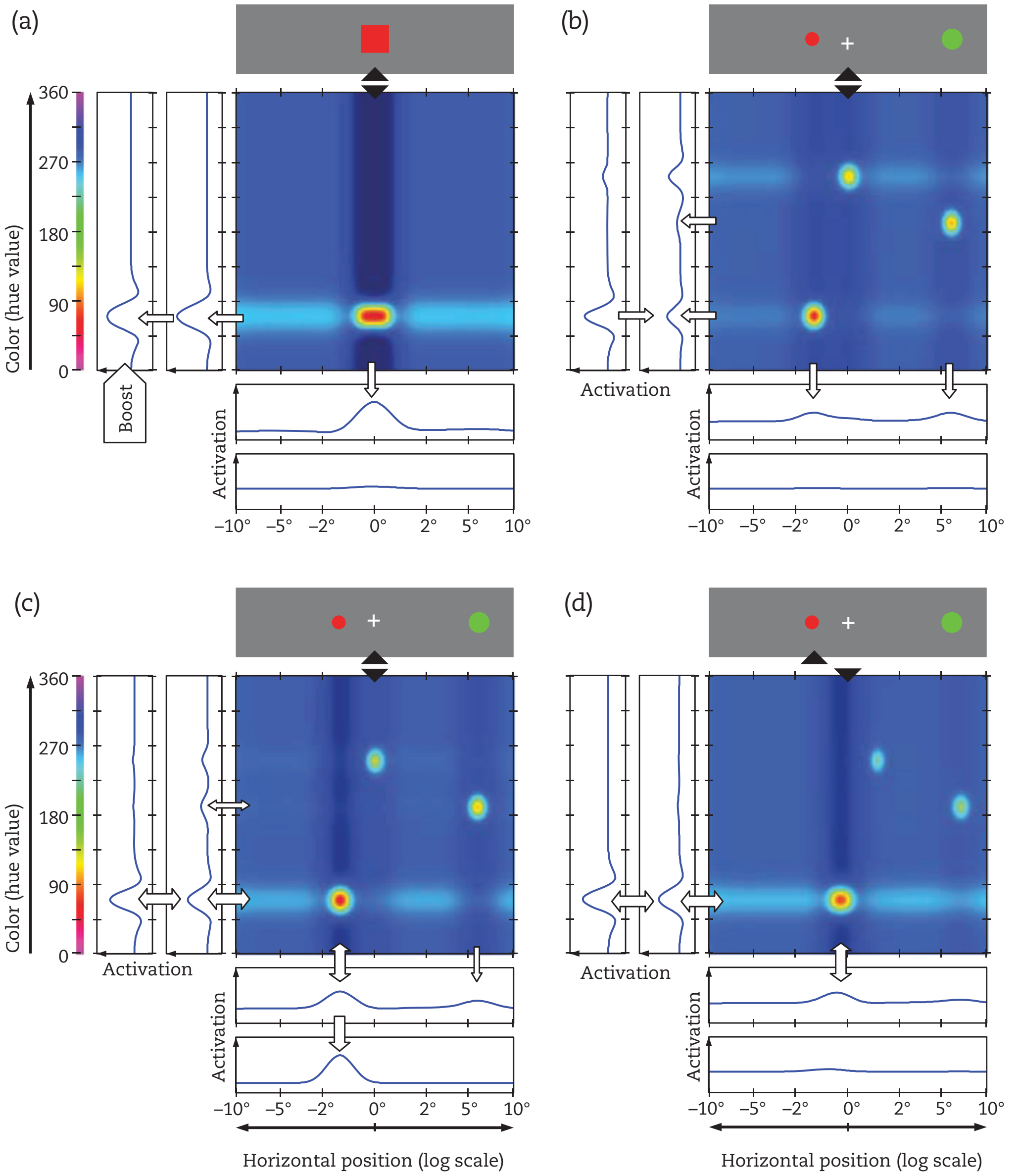

FIGURE 5.9: Simulation of one saccade task trial in the DF model. (*a*) The memory cue is presented in the center of the visual scene, forming a peak in the visual sensory, color attention, and spatial attention field. The color attention field projects activation further to the color memory field, which is boosted during this task period to form a self-sustained peak for the shown color. The saccade motor field is not activated, since the central portion of the spatial attention does not project to it. (*b*) The target and distractor stimuli are presented (together with a fixation point). All visual items compete for spatial attention, realized through the lateral interactions in the spatial attention field. This competition is biased indirectly by the color attention: The red distractor item matches the memorized color, and a peak for this color can form more quickly in the color attention field due to converging input from the memory field. This strengthens the representation of the distractor item in the visual sensory field. (*c*) The location of the distractor item is selected in the spatial attention field, and a peak at the same location is induced in the saccade motor field. This initiates a saccade to the distractor stimulus. (*d*) After the saccade, the distractor is fixated and the retinal scene has shifted accordingly. The memorized color is still represented in the color memory field.

After the memory stimulus is turned off, the activation of the memory control neurons is reduced to an intermediate level. At this level, the peak in the color memory field is sustained by the lateral interactions in this field, but no new peaks can form. The sustained activation in the memory field preactivates the corresponding region in the color attention field. This preactivation remains below the output threshold and does not drive strong neural interactions. Consequently, there is no visible effect of the memory activation in the visual sensory field.

Next, the memory stimulus is replaced by a smaller fixation cue at the same position, which excites weak spatial attention to the foveal region. After the delay, the peripheral target and distractor stimuli appear along with the fixation cue (see Figure 5.9b). All three stimuli produce activation peaks in the visual sensory field. The size of each peak is initially determined by the size and position (distance from the fovea) of each stimulus. As activation in the network continues to cycle, however, feedback from the other fields begins to have an influence.

In the spatial attention field, three active regions are formed by the input from the visual sensory field. These compete with each other via lateral interactions, as well as through indirect interactions within the visual sensory field. This competition is additionally biased by a preshaping of the activation profile in the spatial attention field to account for cognitive inputs reflecting the task instructions. In particular, the foveal area supporting fixation and the region where distractors are presented are suppressed, while the spatial range where targets are presented (farther from the fovea) is preactivated. This reflects the task instructions to make a saccade as fast as possible to the far targets and to ignore the distractor.

At the same time, activation peaks form in the color attention field. The competitive interactions in this field alone are relatively weak, so that at least transiently, multiple peaks may coexist. Only through interactions with the spatial pathway, a selection of a single color is achieved at a later time. There is, however, a significant preactivation in this field at the position of the memorized color due to feedback from the color memory field. If the color of one of the stimuli matches the memorized color, a peak at the corresponding location in the color attention field will form more quickly and reach greater strength than peaks at other positions (Figure 5.9b). The representation of the matching stimulus in the visual sensory field is then strengthened by earlier and stronger feedback from the color attention field. This, in turn, biases the activation profile in the spatial attention field.

The competitive interactions in the spatial attention field amplify small differences in the strengths of individual inputs, and the relatively small biasing effect of the color memory match can change which peak prevails in this field. For the given task, the weak fixation activation is virtually always overcome and either the target or the distractor location is selected. The target stimulus has a competitive advantage due to its larger size and the preactivation of the spatial region in which the targets appear, but this advantage is frequently overcome by the influence of the distractor stimulus if the latter matches the memorized color (as shown in Figure 5.9c). Once a sufficiently strong localized peak has formed in the spatial attention field, it can drive activation in the saccade motor field beyond the output threshold. When this happens, the strong lateral interactions in the motor field and the coupling to the saccade reset neurons produce a stereotypical activation time course, with a rapid formation of a peak at the position of the input and its extinction after a fixed duration. With this mechanism, the amplitude of the eye movement is determined primarily by the location of the activation peak, with only small effects of input strength.

The resulting distribution of saccade amplitudes is shown in Figure 5.8. The distribution is bimodal in all conditions, with one mode centered approximately over the distractor position and one over the range of target positions. The proportion of saccades made to the distractor in each condition reflects the biasing effect of color memory described earlier: It is largest if the distractor stimulus matches the color of the memory stimulus, is reduced in the no-match conditions, and nearly disappears in the target match condition. This pattern accurately reflects the empirical effects. Note that the model also captures saccade latencies from this paradigm, as well as behavioral results from several other variants of the remote distractor paradigm. Interested readers are referred to Schneegans et al. (2014) for details.

MOVING BEYOND COLOR: THE FUNCTION AND FALLIBILITY OF VISUAL FEATURE INTEGRATION

We began this chapter with a consideration of higher-dimensional representations in neurophysiology, such as the two-dimensional representation of retinal space. We discussed how many neural representations in early visual cortex are three-dimensional with a two-dimensional representation of retinal space combined with a single feature such as hue or orientation. We then showed how a simplified DF architecture (designed to capture only changes in the horizontal position of an object) that combined two one-dimensional fields with one two-dimensional field could capture both the integration of features and space as well as the selection of particular features from this integrated representation. Moreover, when we expanded this picture to include a saccade motor system and a feature memory field, we took a first step toward an embodied view of forming a perceptual goal, finding a feature-matching object in the retinal frame, and generating a saccade to this object. This expanded "biased competition" architecture effectively captured behavioral findings from the remote distractor paradigm, showing how DFT can provide a useful framework for thinking about the neural processes that underlie behavior.

In this final section, we push our discussion of higher-dimensional representations one step further. Thus far, we have only addressed how the brain deals with a single non-spatial feature such as color. Of course, we don't just perceive colors; there are neural representations of orientation, spatial frequency, and other complex features involved in, for instance, perceiving a face. Given this, should we just expand our framework to move from two-dimensional fields to three-dimensional and four-dimensional fields and beyond?

A first answer is "no," because this quickly becomes computationally intractable for our computer—and for the nervous system. Recall that a single five-dimensional field sampled using 100 neurons per dimension would require 10 billion (100^5) neurons—compared to just 500 neurons to sample each dimension individually. A second answer is also "no," because the visual cortex does not appear to adopt this approach. Rather, data suggest that there is a division of labor in early visual cortex, with one population of neurons responsive to hue and retinal position, a separate population responsive to orientation and retinal position, a separate population responsive to the direction of motion, and so on (Hubel & Wiesel, 1968, 1977; Livingstone & Hubel, 1984a; Lu & Roe, 2008; Shmuel & Grinvald, 1996).

But this division of labor leads to a critical question: If different populations of neurons represent different features, why don't we have an internal or functional sense of this division of labor? Put differently, how do we perceive objects as integrated wholes? One reason is that there is a simple solution to this variant of what has been called the "binding" problem in vision: an integrated—but distributed—representation of objects can be achieved by coupling all of these neural populations together via the spatial reference frame common to each neural representation.

Concretely, this leads to the expanded biased competition model shown in Figure 5.10. For simplicity, we have added only one additional set of feature-specific fields; in principle, however, we could add more feature-specific fields and achieve the same functionality without a dramatic increase in the neural or computational demands. As can be seen in Figure 5.10, we have added a two-dimensional visual sensory field, a feature attention field, and a feature memory field. Let's assume that these newly added fields represent a continuous shape dimension (see, e.g., Drucker & Aguirre, 2009), so that we now have distinct representations of color and shape. We will refer to the individual visual sensory fields as the space-color field and the space-shape field. The field couplings along the feature pathway are largely equivalent to those in the biased competition model. As a crucial point for the functionality of the multifeature architecture, both visual sensory fields are reciprocally coupled with the spatial attention field. This means that the two visual sensory fields are indirectly connected via spatial attention.

How is a single multifeature item represented in this expanded model? The representation of such an item initially consists of a peak in the space-color field and a peak in the space-shape field (Figure 5.10b). Since both features occupy the same location, the two peaks are aligned along the spatial dimension. The projections from the visual sensory fields to the spatial attention field build a peak there as well, indicating the item's location (Figure 5.10e). Peaks also form in both feature attention fields, indicating the individual feature values of the item (Figure 5.10c). In sum, the multifeature item is represented by the distributed—but coupled—pattern of peaks across these fields.

A problem arises, however, when multiple items are presented simultaneously. As discussed previously, the integration of features and space is lost in the one-dimensional fields, but it is often critical to know which features are present at which location. This leads to a feature-space "binding" problem, which we solved previously by forcing the attention fields to be in a single-peak state. In the case of multifeature items, the problem extends to the conjunction of different features: Which features belong to one item and which belong to another?

It turns out that the same attentional constraint—forcing attention to be selective—solves the feature-conjunction problem as well. In particular, when multiple items are presented to the model, multiple peaks form in each visual sensory

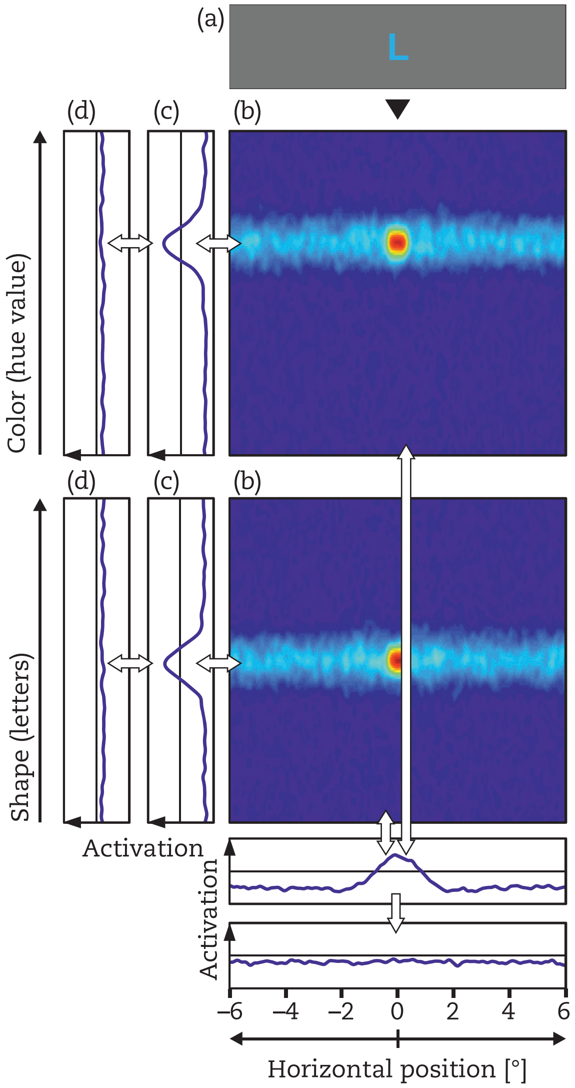

FIGURE 5.10: Architecture of the multifeature model. The model is shown during the presentation of a single item (the blue L). The upper three fields (b, c, d) make up the layer for color, while the bottom three fields (b, c, d) represent the shape layer. (*a*) Visual scene (one-dimensional visual space), illustrating the input to the visual sensory fields. (*b*) Visual sensory fields (*top*: space-color field, *bottom*: space-shape field). (*c*) Feature attention fields. (*d*) Feature memory fields. (*e*) Spatial attention field. (*f*) Spatial "read-out" field (formerly saccade motor field). Note that there are no peaks in the feature memory and the spatial read-out field because these fields require boosts to form peaks.

field—one peak per item. Each of these peaks projects to the spatial attention and feature attention fields. Now, if lateral interaction would allow multiple peaks to coexist in these fields, the resulting pattern of spatial and feature peaks would not include clear information as to which feature belongs to which location or which features belong to the same item. So the attention fields need to be selective and allow only one peak in each field. Through this adjustment, we achieve that when facing multiple items the model by default relaxes into a state where the single peaks in the two feature attention fields belong to the same item. This is because in the spatial attention field as well a single peak develops, and the spatially localized projection of this peak back into the visual sensory fields enhances all peaks at that location—the location of a single item—causing the peaks of that item to project more strongly into the feature attention fields and thus prevail in determining peak position there. What we have then is the coherent representation of one item— but chosen randomly from among all items in the scene.

This solves the problem of mixing up the features of different items, but for the mechanism to be of any practical use a way is needed to control which item is chosen. For example, we might want to retrieve all the features of an item at a specific location. This can be done by forcing the model to build a single peak in the spatial attention field, that is, by having the model focus on a single location.

Assume for the moment that we have some way of selectively "clamping" activation in the field (we will introduce a more natural way later). This will boost the features in the visual sensory fields at that location which, in turn, will selectively create peaks in the feature attention fields.

But what about something more complex—is it possible to attend to color and retrieve shape? Under normal circumstances people do this without any conscious effort, so our model as well should have a way to get from one non-spatial feature to another. In the canonical task used to probe people's ability to do this, participants are briefly shown a multi-item display of multifeature items and are asked to report feature B of the item having a designated target feature A (see, e.g., Ivry & Prinzmetal, 1991). Often, color and letter identity (shape) are used as feature dimensions. The display then consists of a number of colored letters, typically two to five. After brief presentation of the display, the participant has to report, say, the identity of the red letter (or, vice versa, the color of a target letter).

To enable the model to perform this feature integration task, some adjustments to the model parameters are necessary. Note that these changes do not reflect key differences in the way the neural system operates, but rather correspond to changes in its mode of operation—which may adapt to the demands of different tasks. The first specific demand of the task is explicit attention to a stored target feature value (the analogue to what we achieved earlier by "clamping" field activation). Concretely, at the start of each trial, the target-defining feature value is provided by presenting a cue item having this feature. The feature is stored in one of the feature memory fields as a self-sustained peak, triggered by a brief homogeneous boost to the field. Because feature attention to the stored target value should be explicit, we have increased the strength of the projections from the feature memory fields to the feature attention fields, so that a full-fledged peak develops in the feature attention field of the target dimension (instead of only a subthreshold hill of activation). The second novel task demand is the ability to extract a single feature value along the second feature dimension as the feature response. For this, we have made the feature memory fields more selective by increasing inhibition, so that each allows for only one self-sustained peak. A modified version of the task, which we discuss here later, additionally requires participants to indicate not only the second feature but also the spatial location of the target item. For this version a third task demand arises, namely, the ability to extract the spatial location of a single item as the spatial response. For this we use the saccade motor field as a type of spatial "read-out" field that can build a single peak upon receiving a boost, thus indicating one location. This amounts to implementing a simple form of covert spatial attention where we suppress the saccade motor system.

As described earlier, each trial of the task starts with providing the target feature by presenting a cue item and storing one of its features in the target feature memory field. Figure 5.11a shows the phase after presentation of the cue item and before presentation of the test display. In this example, color is the target dimension, and shape is the response dimension. Next, the multi-item test display is shown, letting the model select one item through the dynamic interplay of its components. Decisively, this selection is strongly biased by the color memory peak that causes color attention to be directed at one specific color value (Figure 5.11b), resulting in a ridge of activation in the space-color field. This ridge enhances the space-color peak it overlaps, and the enhanced peak in turn boosts the associated location in the spatial attention field, building a peak there. This peak then projects a spatially localized ridge into the visual sensory fields, enhancing both peaks that belong to the item at the respective location. This includes the space-shape peak of the target item, so that this peak prevails in building a peak in the shape attention field (Figure 5.11c). Lastly, the shape memory field is boosted to obtain the feature response. In the extended version of the task, the spatial read-out field is boosted simultaneously to also obtain an estimate of the item's location, the spatial response.

Taken together, the system is able to determine the shape of an item that was initially selected on the basis of its color. This is achieved through indirect coupling of different feature representations via a shared spatial dimension. This indirect mechanism is consistent with behavioral evidence. In an experiment by Nissen (1985; see also van der Velde & van der Heijden, 1993), people were shown displays of four items, all colored and shaped differently. In the first condition, the participants were cued with a color and had to report the location of the item in that color or vice versa. Probabilities of correct reports were similar for both directions. In the second condition, the participants were cued with a location and had to report the color and the shape of the item in that location. The joint probability of correctly reporting both color and shape

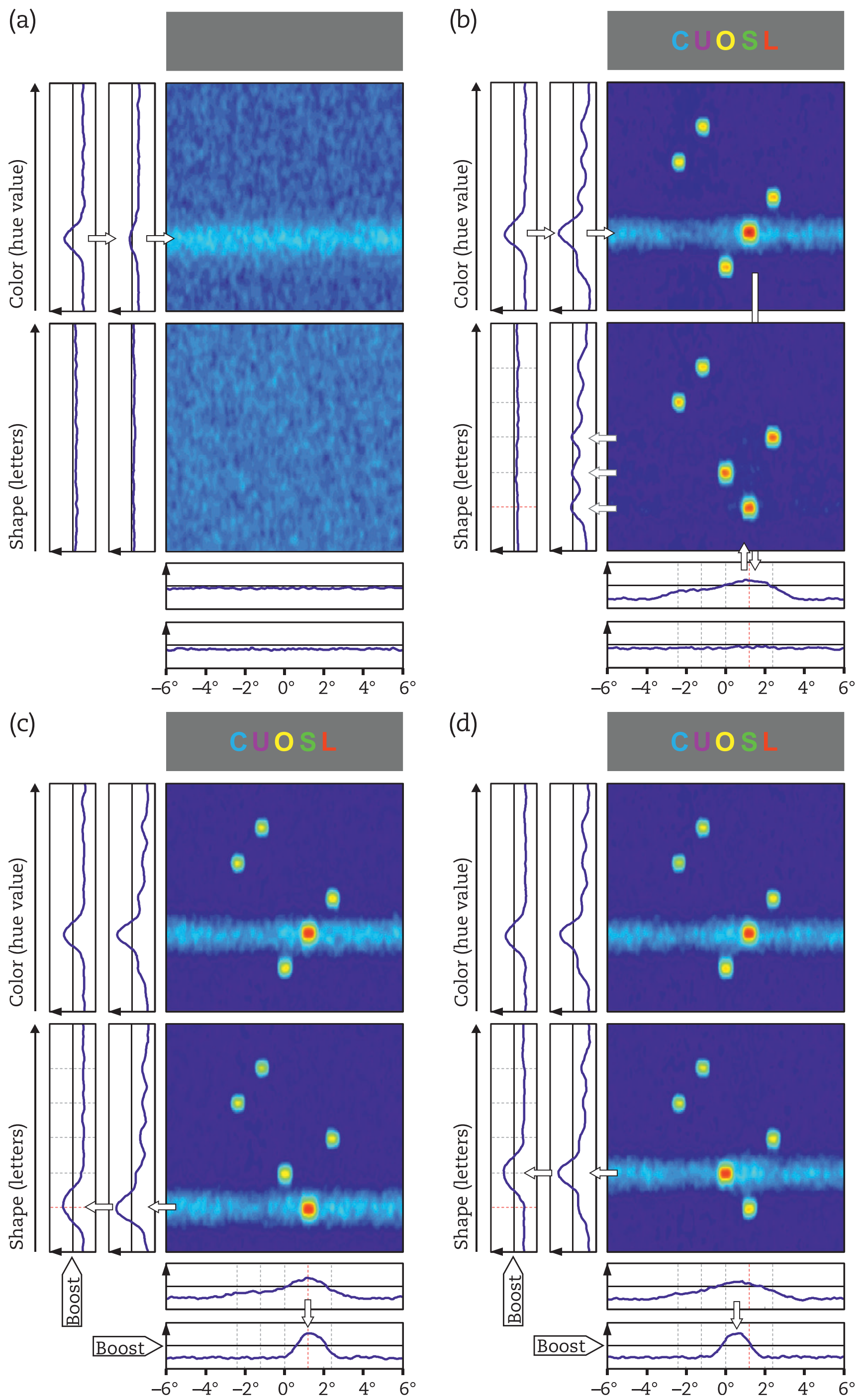

FIGURE 5.11: Multi-item trial in the multifeature model with high spatial proximity and different possible outcomes. (*a*) At the start of each trial, a cue item is presented (not shown) and the color memory field is boosted concurrently. This causes a peak to build there, which is retained throughout the trial and reflects the target color. The projection to the color attention field activates the respective value there, which in turn biases activation in the space-color field. (*b*) Next, the test display with multiple items is presented. Each of the items is represented by one peak in each visual sensory field. The activation ridge from the color attention field enhances the space-color peak of the target item (the green S), causing this peak to determine peak position in the spatial attention field. The spatial attention peak projects back into both visual sensory fields, enhancing the space-shape peak at that location (and less so the peaks of close-by items). (*c*) Brief boosts to the shape memory field and the spatial read-out field force these fields to form peaks, which correspond to the shape and spatial response of the model, respectively. In most cases, the correct shape and location are chosen, as shown here. (*d*) In some cases, the feature-space peak of a distractor item spatially close to the target item (here, the space-shape peak of the yellow O) is overly enhanced by the ridge from the spatial attention field. In this case, the erroneously enhanced peak may prevail in determining peak position in the shape attention field and, thus, the shape response, resulting in an illusory conjunction. Illusory conjunctions are also associated with a shift of peak position in the spatial attention field, which is why the location response is as well displaced toward the spatial midpoint between the involved items.

was equal to the probability of correctly reporting color multiplied by the probability of correctly reporting shape. This suggests that shape and color selection are independent of each other. In the third condition, a color was cued and the participants had to report the shape and the location of the item in that color. Here, the joint probability of correctly reporting both shape and location differed from the product of the individual probabilities, suggesting that selection of one depends on selection of the other. Nissen (1985) hypothesized that the selection of location mediates the selection of shape, rather than vice versa. A testable prediction of this hypothesis is that the joint probability of correctly reporting both shape and location given color (known from the third condition) should be equal to the probability of correctly reporting location given color (known from the first condition) multiplied by the probability of correctly reporting shape given location (known from the second condition). The reason is that once a location has been selected, it should not matter for the selection of shape whether the location was cued directly or selected via color. The data confirmed this prediction, strongly suggesting that the extraction of surface features depends on selecting a location first. Thus, space takes on the role of binding together different surface features.

So the generalized biased competition model can effectively represent multifeature items and extract the location and features of each of them via a mechanism that parallels behavioral findings—but it is not the case that the model performs perfectly under all conditions. Critically, humans aren't perfect, either. And as we discuss here later, the DF model "fails" in precisely the same way humans do.

Although the tasks just discussed sound straightforward, they are actually quite difficult, because in the typical lab tasks the items are flashed very briefly (e.g., for 75 ms) and in the visual periphery. Consequently, people make errors—and the nature of these errors can be informative about how the visual system works. First, there are the so-called feature errors, meaning that sometimes features are reported that were not present in the display. More interestingly, however, subjects are also prone to erroneously combine features from different items, forming so-called illusory conjunctions (ICs; Treisman & Schmidt, 1982). For instance, if a display contains a green X and a red T, observers sometimes report having seen a red X. ICs are thought to reflect a failure at the stage of feature integration rather than at the stage of feature perception itself, because the involved features are actually present in the display but not combined properly by the participant. We will comment more on this interpretation later.

Several factors have been identified to affect the probability that ICs occur. Two key factors are the spatial distance between items (Ashby, Prinzmetal, Ivry, & Maddox, 1996; Cohen & Ivry, 1989) and the similarity of items along the response feature dimension (Ivry & Prinzmetal, 1991). The probability that ICs occur is higher for smaller inter-item distance and higher for more similar feature values (for instance, red versus orange as opposed to red versus green). If our model reflects the neural processes that underlie visual feature integration, it should sometimes make ICs. For instance, the model should sometimes select a shape that does not belong to the red target item but, instead, belongs to one of the distractor items. Such errors indeed occur as we push the model to its limits.

Let us first consider the effect of spatial proximity, using our former example where color is the target-defining dimension and shape is the response dimension. To get integration errors in some trials, we chose a sufficiently small spatial distance between the items in the test display. Together with the coarseness of the selection ridges, this can give rise to a cascade of events that ultimately results in the selection of an incorrect feature. Spatially close peaks in the space-color field project to neighboring regions in the spatial attention field. This can lead to a fused peak in the spatial attention field (Figure 5.11d). This peak is broad and situated somewhere in between the locations of the involved items. Consequently, the ridges sent back into the visual sensory fields are broad and displaced as well. This increases the likelihood that the ridge activation may "spill over" to the visual sensory peaks of the distractor. When this occurs, the position of the spatial ridge may further shift toward the distractor due to the reciprocal coupling with the spatial attention field. The more the distractor item is boosted, the greater the likelihood that the shape attention field will select the wrong shape and make an IC. In this case, the resulting pattern of peaks is composed of the target item's color and the shape of a non-target item (see Figure 5.11d).

The probability for the model to make an IC under these conditions depends mainly on the width of lateral interaction in the spatial attention field, the width of excitatory input from the visual sensory fields to the spatial attention field, and the width of excitatory (ridge) input from the spatial attention field to the visual sensory fields. In addition, inter-item distance must be sufficiently small. “Sufficiently small” in this case means small enough for a spatial ridge to overlap with more than one peak in the visual sensory fields and/or small enough for the projection of two space-color peaks to result in only one spatial attention peak. Thus, as found psychophysically, IC probability in the model is higher for smaller inter-item distance.

Interestingly, we can push this one step further by examining the spatial representation in the model in greater detail by employing the above-mentioned extended version of the task, that is, use the former saccade motor field to “read out” the location of the spatial attention peak. Again, we find systematic errors that mimic human performance. Hazeltine, Prinzmetal, and Elliott (1997) examined where human observers perceive ICs relative to the target and distractor items. They briefly presented strings of five colored letters to their subjects in which one letter was always green. Subsequently, subjects indicated whether the green letter had been an O or not by responding “yes” or “no,” respectively. After that, participants indicated the location on the screen at which they had perceived the green letter. Hazeltine and colleagues were primarily interested in the trials where ICs occurred. They found that in these cases, participants indicated that the green letter was located at the spatial average of the green and O feature locations.

A similar pattern is found in the model. Over trials in which ICs are produced, peak location in the spatial read-out field tends to form a bell-shaped distribution. For small item distances comparable to those used by Hazeltine and colleagues, the center of the distribution tends to lie around the midpoint between the involved items. This makes sense, given that ICs are closely associated with the spatial attention peak being broadened or shifted into the direction of the distracter item. The broader the peak is, or the more it is shifted toward the midpoint between two items, the more likely it is that a trial will yield an IC, and the more likely that the spatial response peak will be centered over the average attended location.

The effect of feature similarity on IC probability can also be observed in the model. ICs occur more often when the target item’s feature value along the response dimension (shape in our example) is close to that of a distractor item. There are two main, interrelated causes for this, which are similar to the effects observed in the spatial case. First, broad projections from the space-shape field to the shape attention field may enhance a distracter peak there that is close to the target item’s peak. Second, an equivalent effect may occur in the converse projection from the shape attention to the space-shape field and, as usually occurs in such cases, the two effects may amplify each other. Statistically, a slightly enhanced distracter peak will determine the final response more often than it would otherwise. For very similar items, a third possibility is that two feature attention peaks fuse into one peak, leading to a deviation of the feature response and possibly resulting in an IC. As is obvious from this description, the two causes for ICs, spatial proximity and feature similarity, are very closely related and probably modulate each other, further elevating IC probability for certain stimulus configurations. Moreover, there may be other causes for ICs, which we are still exploring in detail, such as item similarity along the target dimension.

In summary, the DF model shows several key effects reported in the literature—that ICs are more frequent when items are close together, and when items share similar features on the response dimension. More notably, the model shows the observed spatial averaging effects which have not been effectively explained by other theoretical accounts (e.g., Ashby et al., 1996; Treisman & Gelade, 1980). These parallels with the behavioral literature suggest that the DF model offers a robust theory of visual feature integration and lend support to the most significant conceptual aspect of the model—the role of a shared spatial frame as the basis for visual feature integration.

Note that the effects of spatial proximity seen in our model are similar to those predicted by Ashby and colleagues’ (1996) account, location uncertainty theory (LUT). However, the two models differ in some key respects. In LUT, effects arise due to uncertainty in the position of stimuli during the “feature registration” phase. Similar forms of positional uncertainty could be introduced in our model in various ways, for example, by misplacing inputs in the visual sensory fields. However, this is not necessary to generate ICs, as our model places the main cause for ICs at the level of the attentional mechanisms involved in feature integration, rather than at the level of “feature registration.” LUT does not specify a mechanism for feature integration; rather, it specifies a rule for combining features. The DF model, by contrast, specifies a neural mechanism of feature integration based on a robust framework for thinking about neural population dynamics in early visual processing. Thus, in our view, the DF model offers innovation on several fronts. Nevertheless, given that there is substantial overlap between our theory of visual feature integration and LUT, these two approaches should be viewed as complementary rather than contradictory.

CONCLUSIONS

In this chapter, we formalized a new concept—multidimensional dynamic fields. These fields enable the fast and flexible *integration* of multiple types of information simultaneously. This is critical in early visual processing, and multidimensional representations are pervasive in early visual cortex. Although multidimensional fields are powerful, they are also costly. Moreover, some behaviors do not require full, high-dimensional detail. Thus, we explored the notion of *selection*, using lower-dimensional fields. We grounded this discussion in a concrete example, showing how a DF model of "biased competition" could capture findings from the remote distractor paradigm.

Next, we re-examined how multiple features are represented in early visual cortex. Here, we saw that early visual processing consists of multiple, multidimensional representations, leading to a fundamental question: If different populations of neurons represent different features, why do we perceive items as integrated wholes? We then expanded the biased competition model, adding a second layer of feature-based fields. And, critically, we coupled these fields along a shared spatial dimension. We showed how this can capture both visual-feature integration and the failure of integration in the form of illusory conjunctions.

Although we focused primarily on issues central to early visual processing in this chapter, we also introduced concepts along the way that will be expanded in subsequent chapters. In particular, we used a simplified view of VWM here. The next chapter will pick up on this theme and expand our understanding of this critical form of "active" memory. Similarly, we only focused on a single spatial frame here—the frame of the retina. This is not sufficient; we need to understand how humans anchor decisions, working memories, and memory traces to other spatial reference frames, such as the frame of the head, body, and world. We'll tackle that topic in Chapter 7. And then we'll pull these threads together in Chapters 8 and 9 when we revisit the "binding" problem in vision, but now at the level of a visual scene. We also introduced several modulatory mechanisms in this chapter, such as the memory control nodes. These hint at forms of "higher-level" or "top-down" control. We'll pick up on this topic in several of the later chapters, including Chapters 10 and 14, which examine different senses of cognitive control and executive function.

But also the core integration mechanism with two-dimensional fields that we introduced in this chapter can be employed far beyond the level of early visual processing in problems of "higher-level" cognition. For instance, in Chapter 14, we describe a DF model of executive function that captures people's performance in a card-sorting task. In this task, participants are instructed to sort cards to different specified locations, based on the features of the card. To solve the task, participants have to form an association between each card color and the location where it is to be placed.

The model architecture for this simple association task is the same as that for the basic space-feature integration mechanism, with two one-dimensional fields coupled to a single two-dimensional field. The only difference is that now the two-dimensional field is augmented with a memory trace (see Chapter 2). The task instruction can be given by providing unique inputs to each one-dimensional field. This effectively teaches the system a particular association, for instance, "Put the red card to the left." Each peak in the one-dimensional fields will create an associated ridge in the two-dimensional field. Where these ridges intersect, a peak can form, which in turn lays down a memory trace. The system has now learned that red things go to the left. Similarly, it can learn that green things are to be placed to the right.

After the instruction phase, the DF model can be shown a test card. Here, the color on the card's face is input into the one-dimensional color field, forming an activation peak. This peak, in turn, projects a ridge into the two-dimensional field. If the input ridge overlaps with a localized memory trace, these inputs combine, the output threshold is pierced, and neural interactions combine to form a peak. This peak projects activation into the one-dimensional spatial field and the model "responds" that the card should be placed at the associated spatial location. Note that here, the memory trace provides the localized input for the two-dimensional field, thereby taking the role that was filled by direct visual input in the models described earlier in this chapter. This simple change turns a mechanism for visual search into a system for color-space association based on learned rules.

Before concluding, we want to highlight a few "big picture" ideas that emerge from the issues tackled in this chapter. First, this chapter highlighted several cases where DFT offers a robust view of the integration of brain and behavior. Interestingly, however, this integration was not done in a reductionistic sense. That is, we did not model the brain at one level of analysis and behavior at a different level. Rather, our contention is that neural population dynamics are, ultimately, about behaviors—cognitive behaviors. And, as we showed in Chapter 4, these can be seamlessly interwoven with the sensory and motor surfaces. We'll have more to say about this later. For now, we stress our commitment to both neural and behavioral constraints in the development of theory.

Second, we want to comment on our use of data from several canonical "information-processing" tasks. To place this comment in context, some readers will know that in the 1990s, when dynamic systems theory was first becoming a mainstream theory of perception, action, and cognition, several researchers argued convincingly that dynamic system theory and information processing were incompatible views of mind (see, e.g., Port & van Gelder, 1995; van Gelder, 1998). We resonate with those arguments. And therein lies a source of tension—how can we develop a dynamic systems view of brain and behavior on the one hand, and then spend so much time and energy on findings that have emerged from an incompatible theoretical framework on the other?

We have approached this tension from a very practical perspective. In our view, data generated from an information processing perspective are often robust, puzzling, and demanding of explanation. Moreover, neurophysiologists often adopt an information-processing perspective. Consequently, some of the richest data sets in neuroscience stem from this theoretical vantage point. Nevertheless, there can be a lack of conceptual clarity when we take a dynamic systems approach and capture findings from information-processing tasks. These problems can be compounded when we label DFs with terms like *perception, attention,* and *memory*. Thus, we want to highlight a few key places where our account diverges from information-processing-style thinking.

A common assumption in information-processing accounts is that cognitive and neural systems have dedicated forms of processing—in one variant, such dedicated processing units are called *modules*. Locally, there is a sense of this in DFT in that each DF encodes specific types of information. This encoding provides the local meaning of each field established through the space code principle. Critically, however, this does not mean that "processing" occurs in isolated modules. In the biased competition model, for instance, fields were reciprocally coupled to other neighboring fields. Concretely, this means that the system is not decomposable: the biased competition model is just one large differential equation—a single coupled dynamical system. It is certainly the case that there is structure in this system—the shape attention field doesn't "talk" to the color attention field. But it is not the case that fields can be thought of as isolated processing units.

Similarly, let's consider a second concept that came up in the discussion of illusory conjunctions—the notion of a "stage of processing." DFT certainly retains the idea that neural processing can unfold through a sequence of (neural) events—something akin to a sequence of stages. That said, "stage" means something different in DFT in at least two ways. First, DFT specifies what separates one "stage" of processing from another: the sequences of bifurcations that occur as the system's neural dynamics unfold through time. For instance, input in the biased competition model can build peaks in the visual sensory fields. This can drive a peak in the spatial attention field, which can then drive the selection of peaks for the associated object in the feature attention fields.

Critically, however, the examples in this chapter also highlight that the organization of "stages" (i.e., the sequence of bifurcations) can be flexibly recombined depending on the task and context. For instance, in the IC task, the sequence of bifurcations is very different from what we just described. First, the model must bring up the goal—to detect the target shape—in the memory field. This builds a peak in the attention field. Consequently, when the input comes into the visual sensory fields, the target shape can be boosted. This peak is stronger, which drives a peak in the spatial attention field, which boosts the peak in the color field, which leads to a selection decision in the color attention field.

In summary, then, care must be taken when considering information-processing *concepts* in the context of DFT, even though we will routinely agree with information-processing researchers that their neural and behavioral data are often very cool. Although we see many analogies with information-processing-style of thinking, they are merely that—analogies. In our view, DFT requires a very different mindset. We'll comment on this more in subsequent chapters.

REFERENCES

Ashby, F. G., Prinzmetal, W., Ivry, R., & Maddox, W. T. (1996). A formal theory of feature binding in object perception. *Psychological Review, 103*(1), 165–192.

Blasdel, G. G. (1992). Orientation selectivity, preference, and continuity in monkey striate cortex. *Cortex, 12*(8), 3139–3161.

Blasdel, G. G., & Salama, G. (1986). Voltage-sensitive dyes reveal a modular organization in monkey striate cortex. *Nature, 321*(6070), 579–585.

Cohen, A., & Ivry, R. (1989). Illusory conjunctions inside and outside the focus of attention. *Journal of Experimental Psychology. Human Perception and Performance, 15*(4), 650–663.

Deco, G., & Lee, T. S. (2004). The role of early visual cortex in visual integration: A neural model of recurrent interaction. *European Journal of Neuroscience, 20*(4), 1089–1100.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18*, 193–222.

Desimone, R., & Gross, C. G. (1979). Visual areas in the temporal cortex of the macaque. *Brain Research, 178*(2), 363–380.

Drucker, D. M., & Aguirre, G. K. (2009). Different spatial scales of shape similarity representation in lateral and ventral LOC. *Cerebral Cortex, 19*(10), 2269–2280.

Fazl, A., Grossberg, S., & Mingolla, E. (2009). View-invariant object category learning, recognition, and search: How spatial and object attention are coordinated using surface-based attentional shrouds. *Cognitive Psychology, 58*(1), 1–48.

Hamker, F. H. (2004). A dynamic model of how feature cues guide spatial attention. *Vision Research, 44*(5), 501–521.

Hamker, F. H. (2005). The emergence of attention by population-based inference and its role in distributed processing and cognitive control of vision. *Computer Vision and Image Understanding, 100*(1-2), 64–106.

Hazeltine, R. E., Prinzmetal, W., & Elliott, W. (1997). If it's not there, where is it? Locating illusory conjunctions. *Journal of Experimental Psychology. Human Perception and Performance, 23*(1), 263–277.

Hoffman, J. E., & Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Perception & Psychophysics, 57*(6), 787–795.

Hollingworth, A., Matsukura, M., & Luck, S. L. (2013). Visual working memory modulates rapid eye movements to simple onset targets. *Psychological Science, 24*, 790–796.

Hubel, D. H., & Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *Journal of Physiology, 148*, 574–591.

Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology, 195*(1), 215–243.

Hubel, D. H., & Wiesel, T. N. (1977). Ferrier lecture. Functional architecture of macaque monkey visual cortex. *Proceedings of the Royal Society of London. Series B, 198*(1130), 1–59.

Hübener, M., Shoham, D., Grinvald, A., & Bonhoeffer, T. (1997). Spatial relationships among three columnar systems in cat area 17. *Journal of Neuroscience, 17*(23), 9270–9284.

Irwin, D. (1992). Memory for position and identity across eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(2), 307–317.

Irwin, D., & Gordon, R. D. (1998). Eye movements, attention, and transsaccadic memory. *Visual Cognition, 5*, 127–155.

Issa, N. P., Trepel, C., & Stryker, M. P. (2000). Spatial frequency maps in cat visual cortex. *Journal of Neuroscience, 20*(22), 8504–8514.

Ivry, R. B., & Prinzmetal, W. (1991). Effect of feature similarity on illusory conjunctions. *Perception & Psychophysics, 49*(2), 105–116.

Kopecz, K., & Schöner, G. (1995). Saccadic motor planning by integrating visual information and pre-information on neural dynamic fields. *Biological Cybernetics, 73*(1), 49–60.

Kowler, E., Anderson, E., Dosher, B., & Blaser, E. (1995). The role of attention in the programming of saccades. *Vision Research, 35*(13), 1897–1916.

Livingstone, M. S., & Hubel, D. H. (1984a). Specificity of intrinsic connections in primate primary visual cortex. *Journal of Neuroscience, 4*(11), 2830–2835.

Livingstone, M., & Hubel, D. (1984b). Anatomy and physiology of a color system in the primate visual cortex. *Journal of Neuroscience, 4*(1), 309–356.

Lu, H. D., & Roe, A. W. (2008). Functional organization of color domains in V1 and V2 of macaque monkey revealed by optical imaging. *Cerebral Cortex, 18*(3), 516–533.

Marino, R. A., Trappenberg, T. P., Dorris, M., & Munoz, D. P. (2012). Spatial interactions in the superior colliculus predict saccade behavior in a neural field model. *Journal of Cognitive Neuroscience, 24*(2), 315–336.

Mishkin, M., & Ungerleider, L. G. (1982). Contribution of striate inputs to the visuospatial functions of parieto-preoccipital cortex in monkeys. *Behavioural Brain Research, 6*(1), 57–77.

Nissen, M. J. (1985). Accessing features and objects: Is location special? In M. I. Posner & O. S. M. Marin (Eds.), *Attention and performance XI* (pp. 205–219). Hillsdale, NJ: Erlbaum.

Owen, M. R., & Laing, C. R., & Coombes, S. (2007). Bumps and rings in a two-dimensional neural field: Splitting and rotational instabilities. *New Journal of Physics*, *9*(10), 378.

Port, R. F., & van Gelder, T. (1995). *Mind as motion: Explorations in the dynamics of cognition*. Cambridge, MA: MIT Press.

Schmidt, B. K., Vogel, E. K., Woodman, G. F., & Luck, S. J. (2002). Voluntary and automatic attentional control of visual working memory. *Perception & Psychophysics*, *64*(5), 754–763.

Schneegans, S., Spencer, J. P., Schöner, G., Hwang, S., & Hollingworth, A. (2014). Dynamic interactions between visual working memory and saccade target selection. *Journal of Vision*, *14*(11), 9.

Shmuel, A., & Grinvald, A. (1996). Functional organization for direction of motion and its relationship to orientation maps in cat area 18. *Journal of Neuroscience*, *16*(21), 6945–6964.

Swindale, N. V., Shoham, D., Grinvald, A., Bonhoeffer, T., & Hübener, M. (2000). Visual cortex maps are optimized for uniform coverage. *Nature Neuroscience*, *3*(8), 822–826.

Trappenberg, T. P., Dorris, M. C., Munoz, D. P., & Klein, R. M. (2001). A model of saccade initiation based on the competitive integration of exogenous and endogenous signals in the superior colliculus. *Journal of Cognitive Neuroscience*, *13*(2), 256–271.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136.

Treisman, A., & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology*, *14*(1), 107–141.

van der Velde, F., & van der Heijden, A. H. (1993). An element in the visual field is just a conjunction of attributes: A critique of Bundesen (1991). *Perception & Psychophysics*, *53*(3), 345–349; discussion 350–352.

van Gelder, T. (1998). The dynamical hypothesis in cognitive science. *Behavioral and Brain Sciences*, *21*(5), 615–628; discussion 629–665.

Wachtler, T., Sejnowski, T. J., & Albright, T. D. (2003). Representation of color stimuli in awake macaque primary visual cortex. *Neuron*, *37*(4), 681–691.

Wilimzig, C., Schneider, S., & Schöner, G. (2006). The time course of saccadic decision making: Dynamic field theory. *Neural Networks*, *19*(8), 1059–1074.

EXERCISES FOR CHAPTER 5

Exercise 1: Basic Instabilities in a Two-Dimensional Dynamic Field

Start the two-dimensional field simulator by running the file `launcherOneLayerField2D`. This simulator is very similar to the one-dimensional, one-layer simulator from Chapter 2, except that the field is now defined over a two-dimensional space. The two visualizations in the graphical user interface (GUI) show the field activation and field output (sigmoid function of the field activation), both color coded. The input and the interaction kernel are not shown separately. There are two localized stimuli, s_1 and s_2, for which you can control both x- and y-position, and two ridge inputs, s_3 and s_4, for which only the position along one dimension can be controlled. For an alternative visualization of the activation distribution in the two-dimensional field, you can also run the file `launcherOneLayerField2D _ surfacePlot`, which is functionally equivalent.

Find appropriate interaction strengths c_{exc}, c_{inh}, and c_{glob} to reproduce all of the basic instabilities in this two-dimensional field:

a) *Detection instability*: Formation of a peak from localized input, with bistability and hysteresis effect when the stimulus amplitude is varied (use the visualization of the output to clearly distinguish between supra-threshold peaks and subthreshold hills of activation).
b) *Selection instability*: Formation of a single peak from multiple localized inputs, and suppression of the other inputs.
c) *Memory instability*: Peaks remain self-sustained even if the input is turned off completely; either limited to a single peak when using global inhibition, or allowing multiple memory peaks with surround inhibition.

Exercise 2: Ridge inputs

Explore the behavior of the field with ridge inputs (s_3 and s_4):

a) Return to interaction settings that create a detection instability. Set up one or two localized stimuli that are too weak to induce peaks (only subthreshold hills of activation, with stimulus amplitude lower than the negative resting level, $a_S < -h$). Now use a weak ridge input that overlaps with one of the localized stimuli to induce a supra-threshold activation peak.
b) Choose interaction strengths that produce selection decisions. Use a ridge input to bias a selection between two localized inputs. First set up the stimuli, then use the Reset button to observe the selection.

In both cases, vary the overlap between the ridge input and the localized input. For the selection, also vary the distance between the two peaks in one dimension, so that a single ridge overlaps with both of them.

You can also try to create localized peaks from a ridge input alone. You can get either a single peak or multiple peaks along the ridge, depending on the type and strength of inhibition. Note that the peak location is random along the dimension for which the ridge input is homogenous. Finally, you can create peaks from the intersection of orthogonal ridges.

Exercise 3: Coupling Between One-Dimensional and Two-Dimensional Fields

Run the file `launcherCoupling`. This simulation implements the space-feature association mechanism described in this chapter. The fields in the simulation are identified by one-letter indices: *v* for the two-dimensional visual field, *s* for the one-dimensional spatial field, and *f* for the one-dimensional surface feature field. Sliders exist for setting the lateral interaction strengths for all three fields, as well as the coupling strengths between the fields (hover over each slider to get a description of the affected parameter). In addition, on the right side of the GUI there are sliders to set localized stimuli for all three fields

Set appropriate interaction strengths to reproduce different coupling behaviors described in the chapter:

a) Induce two activation peaks in field *v* by localized inputs, and set up lateral interactions in that field to make the peaks stabilized. Perform a read-out operation into the two separate one-dimensional fields. Set projection strengths such that peaks are induced in both field *s* and field *f*. Add weak feedback projections from the one-dimensional fields to field *v*. If necessary, adjust lateral inhibition strengths to prevent excessive growth of activation levels.

b) Now reduce the coupling strengths such that the input from field *v* induces only subthreshold activation in both one-dimensional fields, but no peaks. Create a selection regime for these two fields by setting the lateral interaction strengths to appropriate values (local excitation and global inhibition, no local inhibition). Now boost field *s* (by increasing its resting level h_s) to randomly select the location of one stimulus, and then boost field *f* to select the surface feature value of the stimulus at that location. Make sure that there is an indirect coupling between the two one-dimensional fields via field *v*, and that field *f* always selects the feature value that belongs to the item spatially selected in field *s*.

c) With the same settings and the same localized inputs in field *v*, perform a visual search for a surface feature. Induce a peak in field *f* by a direct stimulus input to this field. Now select the location of the matching item by boosting field *s*.

d) Implement the memory-based feature-space association for the card-sorting task described in this chapter. Make the localized inputs to field *v* subthreshold (this simulates the memory trace for learned associations). Set the one-dimensional fields up in a regime where they can have multiple peaks. Induce a peak in field *f* by a direct stimulus. This peak should induce a detection decision in field *v* if the resulting ridge input is close enough to one of the localized hills of activation. The spatial location of that peak in *v* should then be read out into the spatial field *s*.