1

Spatiotemporal Dynamics of Visual Perception Across Neural Maps and Pathways

Haluk Öğmen

Department of Electrical and Computer Engineering
Center for Neuro–Engineering and Cognitive Science
University of Houston
Houston, TX 77204–4005 USA
ogmen@uh.edu

1.1 Introduction

The relationship between geometry and brain function presents itself as a dual problem: on the one hand, since the basis of geometry is in brain function, especially that of the visual system, one can ask what the brain function can tell us about the genesis of geometry as an abstract form of human mental activity. On the other hand, one can also ask to what extent geometry can help us understand brain function. Because the nervous system is interfaced to our environment by sensory and motor systems and because geometry has been a useful language in understanding our environment, one might expect some convergence of geometry and brain function at least at the peripheral levels of the nervous system. Historically, there has been a close relationship between geometry and theories of vision starting as early as Euclid. Given light sources and an environment, one can easily calculate the corresponding images on our retinae using basic physics and geometry. This is usually known as the "forward problem" [41]. A straightforward approach would be then to consider the function of the visual system as the computation of the inverse of the transformations leading to image formation. However, this "inverse optics" approach leads to ill-posed problems and necessitates the use of a priori assumptions to reduce the number of possible solutions. The use of a priori assumptions in turn makes the approach unsuitable for environments that violate the assumptions. Thus, the inverse optics formulation fails to capture the robustness of human visual perception in complex environments. On the other hand, visual illusions, i.e. discrepancies between the physical stimuli and the corresponding percepts, constitute examples of the limitations of the human visual system. Nevertheless, these illusions do not affect significantly the overall performance of the system, as most people operate succesfully in the environment without even noticing these illusions. The illusions are usually discovered by scientists, artists, and philosophers who scrutinize deeply the re-

lation between the physical and psychological world. These illusions are often used by vision scientists as "singular points" to study the visual system.

How the inputs from the environment are transformed into our conscious percepts is largely unknown. The goals of this chapter are twofold: first, it provides a brief review of the basic neuroanatomical structure of the visual system in primates. Second, it outlines a theory of how neural maps and pathways can interact in a dynamic system, which operates principally in a transient regime, to generate a spatiotemporal neural representation of visual inputs.

1.2 The Basic Geometry of Neural Representation: Maps and Pathways

The first stage of input representation in the visual system occurs in the retina. The retina is itself a complex structure comprising five main neuronal types organized in direct and lateral structures (Fig. 1). The "direct structure"
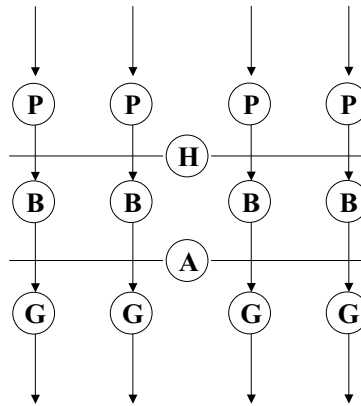

Fig. 1.1. The general architecture of the retina. *P*, photoreceptor; *B*, bipolar cell; *G*, ganglion cell; *H*, horizontal cell; *A*, amacrine cell. The *arrows on top* show the light input coming from adjacent spatial locations in the environment, and the *arrows at the bottom* represent the output of the retina, which preserves the two-dimensional topography of the inputs. This gives rise to "retinotopic maps" at the subsequent processing stages

consists of signal flow from the photoreceptors to bipolar cells, and finally to retinal ganglion cells, whose axons constitute the output of the retina. This direct pathway is repeated over the retina and thus constitutes an "image plane" much like the photodetector array of a digital camera. In addition to

the cells in the direct pathway, horizontal and amacrine cells carry out signals laterally and contribute to the spatiotemporal processing of the signals. Overall, the three-dimensional world is projected to a two-dimensional *retinotopic map* through the optics of the eye, the two-dimensional sampling by the receptors, and the spatial organization of the post-receptor direct pathway. The parallel fibres from the retina running to the visual cortex via the lateral geniculate nucleus (LGN) preserve the retinal topography, and the early visual representation in the visual cortex maintains the retinotopic map.

In addition to this spatial coding, retinal ganglion cells can be broadly classified into three types: P, M, and K [15, 27]. The characterization of the K type is not fully detailed, and our discussion will focus on the M and P types. These two cell types can be distinguished on the basis of their anatomical and response characteristics; for example, M cell responses have shorter latencies and are more transient than P cell responses [16, 33, 36, 42]. Thus the information from the retina is not carried out by a single retinotopic map, but by three maps that form *parallel pathways*. Moreover, different kinds of information are carried out along these pathways. The pathway originating from P cells is called the parvocellular pathway, and the pathway originating from M cells is called the magnocellular pathway.

The signals that reach the cortex are also channeled into maps and pathways. Two major cortical pathways, the dorsal and the ventral, have been identified (Fig. 1.2) [35]. The dorsal pathway, also called the "where pathway", is specialized in processing information about the position of objects. On the other hand, the ventral pathway, also called the "what pathway", has been implicated in the processing of object identities [35]. Another related functional interpretation of these pathways is that the dorsal pathway is specialized for action, while the ventral pathway is specialized for perception [34]. This broad functional specialization is supplemented by more specialized pathways dedicated to the processing of motion, color, and form [32, 59]. Within these pathways, the cortical organization contains maps of different object attributes. For example, neurons in the primary visual cortex respond preferentially to the orientations of edges. Spatially, neurons that are sensitive to adjacent orientations tend to be located in adjacent locations forming a "map of orientation" on the cortical space [30]. This is shown schematically in Fig. 1.3. Similar maps have been observed for location (retinotopic map) [30], spatial frequency [19], color [52, 58], and direction of motion [2].

Maps build a relatively continuous and periodic topographical representation of stimulus properties (e.g., spatial location, orientation, color) on cortical space. What is the goal of such a representation? In neural computation, in addition to the processing at each neuron, a significant amount of processing takes place at the synapses. Because synapses represent points of connection between neurons, functionally both the development and the processing characteristics of the synapses are often specialized based on processing and encoding characteristics of both pre- and post-synaptic cells. Consequently, map representations in the nervous system appear to be correlated with the
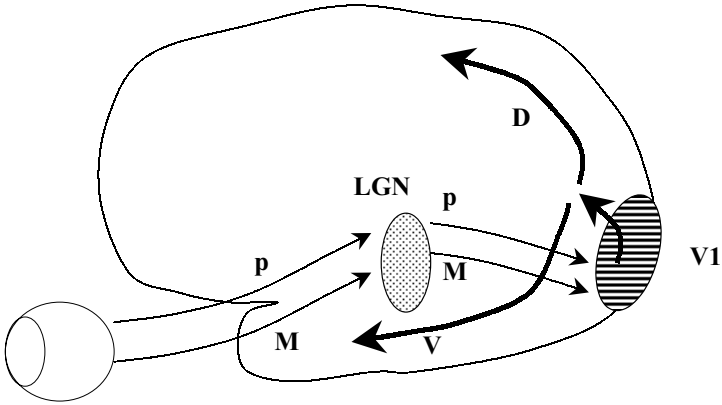
Fig. 1.2. Schematic depiction of the parvocellular (*P*), magnocellular (*M*), and the cortical dorsal (*D*), ventral (*V*) pathways. *LGN*, lateral geniculate nucleus; *V1*, primary visual cortex
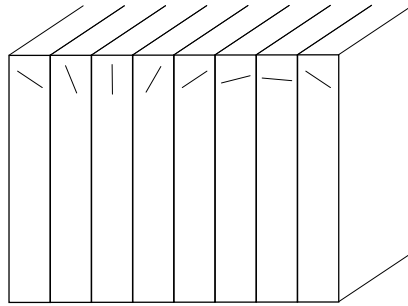

Fig. 1.3. Depiction of how orientation columns form an orientation map. Neurons in a given column are tuned to a specific orientation depicted by an *oriented line segment* in the figure. Neurons sensitive to similar orientations occupy neighboring positions on the cortical surface

geometry of synaptic development as well as with the geometry of synaptic patterns as part of information processing. According to this perspective, maps represent the geometry of the fusion between structure and function in the nervous system.

On the other hand, pathways possess more discrete, often dichotomic, representation. What is more important, pathways represent a cascade of maps that share common functional properties. From the functional point of view, pathways can be viewed as complementary systems adapted to conflicting but complementary aspects of information processing. For example, the magnocellular pathway is specialized for processing high-temporal low-spatial frequency information, whereas the parvocellular system is specialized for processing low-temporal and high-spatial frequency information. From the evolutionary point of view, pathways can be viewed as new systems that emerge as the interactions between the organism and the environment become more sophisticated. For example, for a simple organism the localization of stimuli without complex recognition of its figural properties can be sufficient for survival. Thus a basic pathway akin to the primate where/action pathway would be sufficient. On the other hand, more evolved animals may need to recognize and categorize complex aspects of stimuli, and thus an additional pathway specialized for conscious perception may develop.

In the next section, these concepts will be illustrated by considering how the visual system can encode object boundaries in real-time.

1.3 Example: Maps and Pathways in Coding Object Boundaries

1.3.1 The Problem of Boundary Encoding

Under visual fixation conditions, the retinal image of an object boundary is affected by the physical properties of light, the optics of the human eye, the neurons and blood vessels in the eye, eye movements, and the dynamics of the accommodation system [19]. Several studies show that processing time on the order of 100 ms is required in order to reach "optimal" form and sharpness discrimination [4, 11, 29, 55] as well as more veridical perception of the sharpness of edges [44].

A boundary consists of a change of a stimulus attribute, typically luminance, over space. Because this change can occur rapidly for sharp boundaries and gradually for blurred boundaries, measurements at multiple scales are needed to detect and code boundaries and their spatial profile. The visual system contains neurons that respond preferentially to different spatial frequency bands. Moreover, as mentioned in the previous section, these neurons are organized as a "spatial frequency map" [19, 51]. The rate of change of a boundary's spatial profile also depends on the contrast of the boundary as shown in Fig. 1.4. For a fixed boundary transition width (e.g. w_1 in Fig.
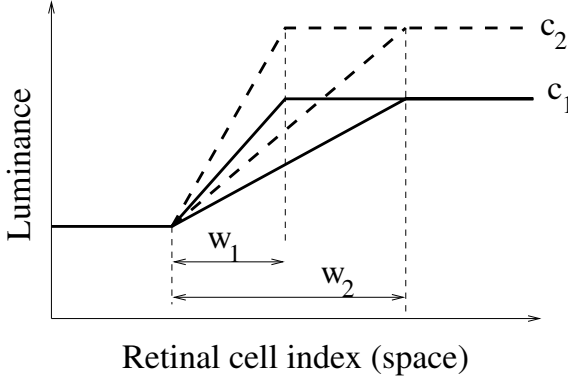
Fig. 1.4. The relationship between contrast and blur for boundaries. Boundary transition widths w_1 and w_2 for boundaries at a low contrast level c_1 (*solid lines*) and a high contrast level c_2 (*dashed lines*)

1.4), the slope of the boundary increases with increasing contrast (c_1 to c_2 in Fig. 1.4). The human visual system is capable of disambiguating the effects of blur and contrast, thereby generating conrast-independent perception of blur [23]. On the other hand, *discrimination* of edge blur depends on contrast, suggesting that the visual system encodes the blur of boundaries at least at two levels, one of which is contrast dependent, and one of which is contrast independent.

1.3.2 A Theory of Visual Boundary Encoding

How does the visual system encode object boundaries and edge blur in real-time? We will present a model of retino-cortical dynamics (RECOD) [37, 44] to suggest (i) how maps can be used to encode the position, blur, and contrast of boundaries; and (ii) how pathways can be used to overcome the real-time dynamic processing limitations of encoding across the maps. The fundamental equations of the model and their neurophysiological bases are given in the Appendix. Detailed and specialized equations of the model can be found in [44].

Figure 1.5 shows a diagrammatic representation of the general structure of RECOD. The lower two populations of neurons correspond to retinal ganglion cells with slow-sustained (parvo) and fast-transient (magno) response properties [16, 33, 36, 42]. Each of these populations contains cells sampling different retinal positions and thus contains a spatial (retinotopic) map. Two pathways, parvocellular (P pathway) and magnocellular (M pathway), emerge from these populations. These pathways provide inputs to post-retinal areas. The model also contains reciprocal inhibitory connections between post-retinal areas that receive their main inputs from P and M pathways. Figure 1.6 shows a more detailed depiction of the model. Here, circular symbols depict neurons whose
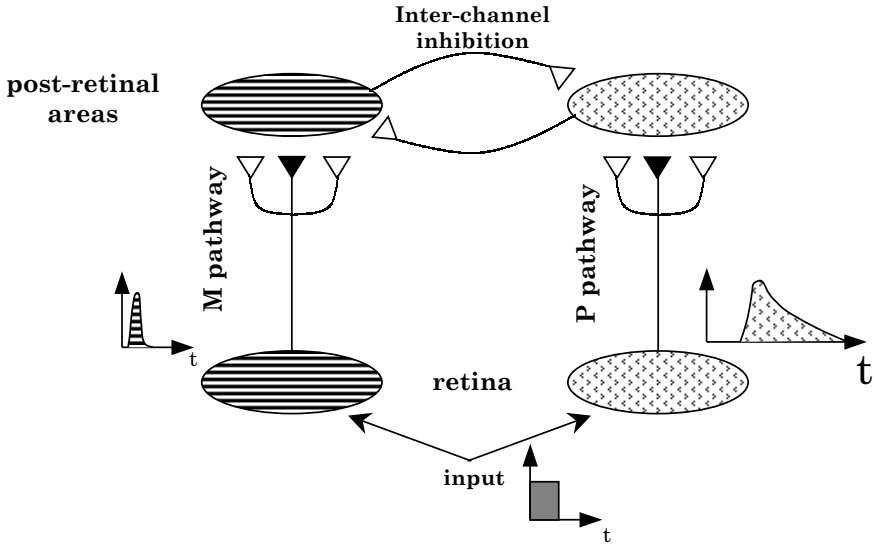
Fig. 1.5. Schematic representation of the major pathways in the RECOD model. *Filled and open synaptic symbols* depict excitatory and inhibitory connections, respectively

spatial relationship follows a retinotopic map. In this figure, the post-retinal area that receives its major input from the P pathway is decomposed into two layers. Both layers preserve the retinotopic map and add a spatial-frequency map (composed of the spatial-frequency channels). For simplicity, only three elements of the spatial-frequency map ranging from the highest spatial frequency class (H) to the lowest spatial frequency class (L) are shown. The M pathway sends a retinotopically organized inhibitory signal to cells in the first post-retinal layer. The direct inhibitory connection from retinal transient cells to post-retinal layers is only for illustrative purpose; in vivo the actual connections are carried out by local inhibitory networks. The first post-retinal layer cells receive center-surround connections from the sustained cells (parvocellular pathway). The rows indicated by H, M, and L represent elements with high, medium, and low spatial frequency tuning in the spatial frequency map, respectively. Each of the H, M, and L rows in the first post-retinal layer receive independent connections from the retinal cells, and there are no interactions between the rows. Cells in the second post-retinal layer receive center-surround connections from the H, M, and L rows of the first post-retinal layer. They also receive center-surround feedback. Sample responses of model
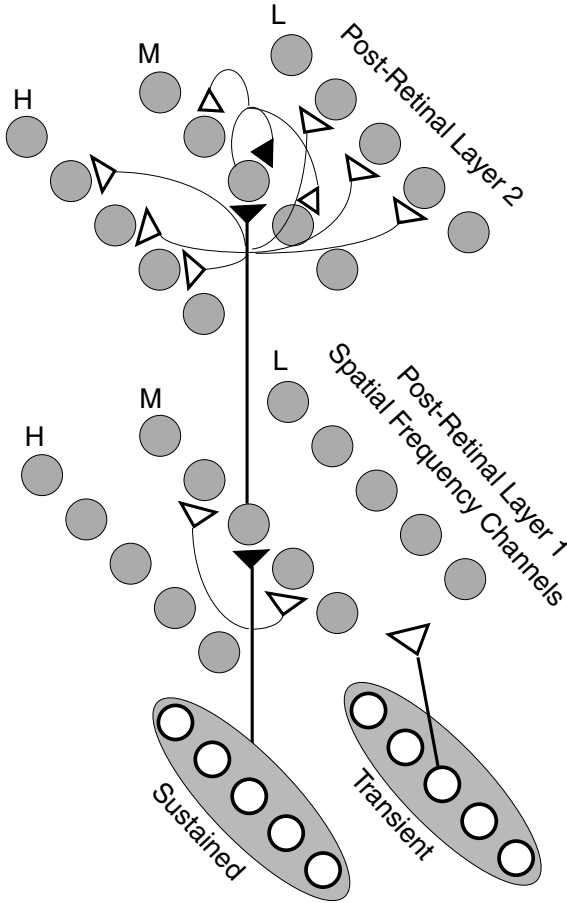
Fig. 1.6. A more detailed depiction of the RECOD model. *Filled and open synaptic symbols* depict excitatory and inhibitory connections, respectively. To avoid clutter, only a representative set of neurons and connections are shown. From [44]

neurons tuned to low spatial frequencies and to high spatial frequencies are shown for sharp and blurred edge stimuli in Fig. 1.7. As one can see in the left panel of this figure, for a sharp edge neurons in the high spatial-frequency channel respond more strongly (dashed curve) compared to neurons in the low spatial-frequency channel (solid curve). Moreover, neurons tuned to low spatial-frequencies tend to blur sharp edges. This can be seen by comparing the spread of activity shown by the dashed and solid curves in the left panel. The right panel of the figure shows the responses of these two channels to a blurred edge. In this case, neurons in the low spatial-frequency channel respond more strongly (solid curve) compared to neurons in the high spatial-frequency channel. Overall, the peak of activity across the spatial-frequency
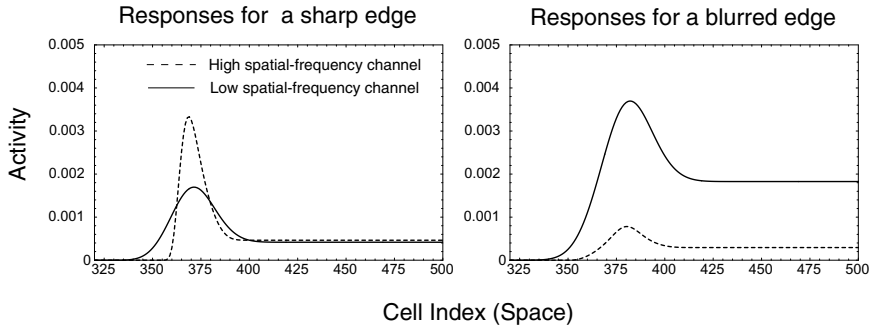
Fig. 1.7. Effect of edge blur on model responses: model responses in the first post-retinal layer for sharp (*left*) and blurred (*right*) edges at high spatial-frequency (*dotted line*) and low spatial-frequency (*continuous line*) loci of the spatial-frequency map. From [44]

map will indicate which neuron's spatial frequency matches best the sharpness of the input edge, and the level of activity for each neuron for a given edge will provide a measure of the level of match. Thus the *distribution* of activity across the spatial-frequency map provides a measure of edge blur. Even though the map is discrete in the sense that it contains a finite set of neurons, the distribution of activity in the map can provide the basis for a fine discrimination and perception of edge blur. This is similar to the encoding of color, where the distributed activities of only three primary components provide the basis for a fine discrimination and perception of color.

The model achieves the spatial-frequency selectivity by the strength and spatial distribution of synaptic connections from the retinal network to the first layer of the post-retinal network. A neuron tuned to high spatial frequencies receives excitatory and inhibitory inputs from a small retinotopic neighborhood, while a neuron tuned to low spatial frequencies receives excitatory and inhibitory inputs from a large retinotopic neighborhood (Fig. 1.8). Thus the retinotopic map allows the simple geometry of neighborhood and the resulting connectivity pattern to give rise to spatial-frequency selectivity. By smoothly changing this connectivity pattern across cortical space, one obtains a spatial-frequency map (e.g. L, M, and H in Fig. 1.6), which in turn, as mentioned above, can relate the geometry of neural activities to the fine coding of edge blur.

The left panel of Fig. 1.9 shows the activities in the first post-retinal layer of the model for a low (dashed curve) and a high (solid curve) contrast input. The response to the high contrast input is stronger. The first post-retinal layer in the model encodes edge blur in a contrast-dependent manner. The second post-retinal layer of cells achieves contrast-independent encoding of edge blur. Contrast independence is produced through connectivity patterns that exploit retinotopic and spatial-frequency maps. The second post-retinal layer implements retinotopic center-surround shunting between the cells in
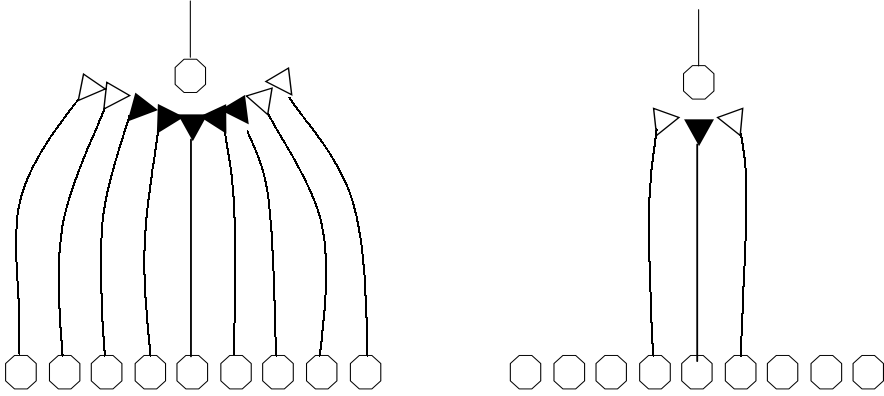
Fig. 1.8. The connectivity pattern on the *left* produces low spatial-frequency selectivity because of the convergence of inputs from an extended retinotopic area. The connectivity pattern on the *right* produces a relatively higher spatial frequency selectivity

the spatial frequency map. Each cell in this layer receives center excitation from the cell at its retinotopic location and only one of the elements in the map below it. However, it receives surround inhibition from all the elements in the map in a retinotopic manner, from a neighborhood of cells around its retinotopic location [12, 18, 20, 49, 50]. In other words, excitation from the bottom layer is one-to-one whereas inhibition is many-to-one pooled activity. This shunting interaction transforms the input activity $p1_i$ for the ith element in the spatial frequency map into an output activity $p2_i = p1_i/(A1 + \sum_i p1_i)$, where $A1$ is the time constant of the response [12, 25]. Therefore, when the total input $\sum_i p1_i$ is large compared to to $A1$, the response of each element in the spatial frequency map is contrast-normalized across the retinotopic map, resulting in contrast-constancy. This is shown in the right panel of Fig. 1.9: the responses to low contrast (dashed curve) and high contrast (solid curve) are identical.

In order to compensate the blurring effects introduced at the retinal level, the RECOD model uses a connectivity pattern across retinotopic maps, but instead of being feedforward as those giving rise to spatial-frequency selectivity, these connections are *feedback* (or re-entrant), as illustrated at the top of Fig. 1.6. Note that, for simplicity, in this figure only the connections for the medium spatial frequencies (M) are shown. Because of these feedback connections and the dynamic properties of the network, the activity pattern is "sharpened" in time to compensate for the early blurring effects. [25, 37]. In
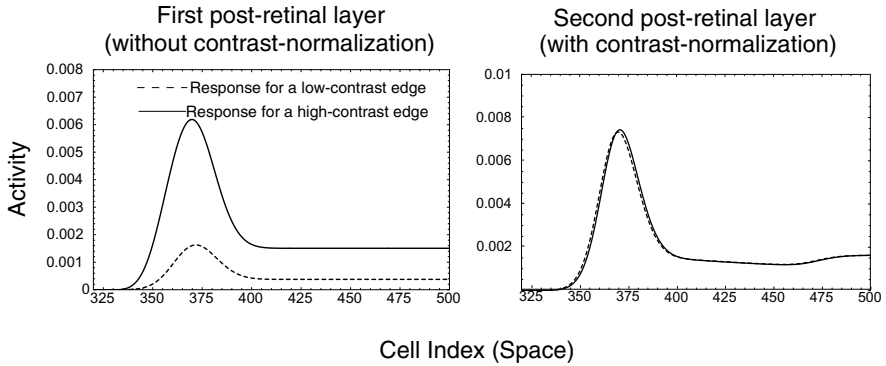
First post-retinal layer
(without contrast-normalization)

Second post-retinal layer
(with contrast-normalization)

Fig. 1.9. Effect of contrast on model responses: Model responses for a high-contrast edge (*solid curve*) and a low-contrast edge (*dashed curve*) of 2 arcmin blur in the first post-retinal layer (*left*) and the second post-retinal layer (*right*). From [44]
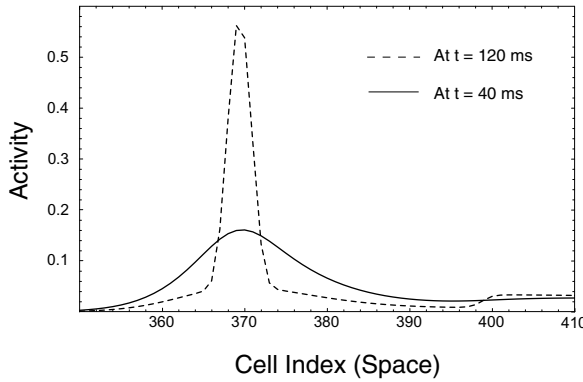

Fig. 1.10. Temporal sharpening of model responses to a blurred edge in the second post-retinal layer: responses at 40 ms (*continuous line*) and 120 ms (*dashed line*) are shown superimposed. From [44]

Fig. 1.10, the response of the model neurons in the second post-retinal layer to an edge stimulus with 2 arcmin base blur at 40 ms after stimulus onset is shown by the dashed curve. The response at 120 ms after stimulus onset is shown by the solid curve. Comparing the width of these activities, one can see that the neural encoding of the edge is initially (at 40 ms) blurred but becomes sharper with more processing time (at 120 ms).

1.3.3 Perception and Discrimination of Edge Blur

The proposed encoding scheme across retinotopic and spatial-frequency maps has been tested by comparing model predictions to a wide range of experimental data [44]. For example, Fig. 1.11 provides a comparison of model

predictions to experimental data on the effect of exposure duration on perceived blur for base blurs of 0, 2, and 4 arcmin. The model has been also
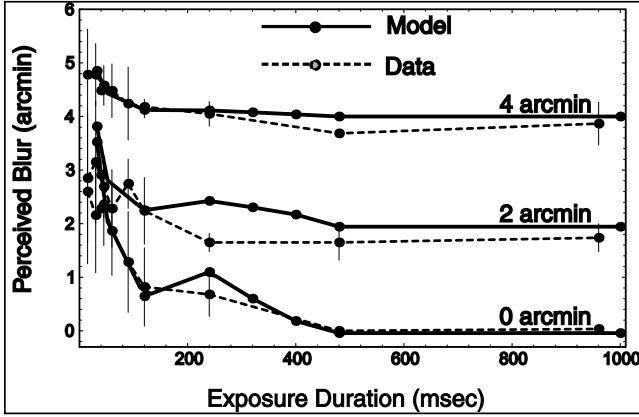

Fig. 1.11. Model predictions (*solid lines*) and data (*dashed lines*) for the effect of exposure duration on perceived blur for base blurs of 0, 2, and 4 arcmin. From [44]
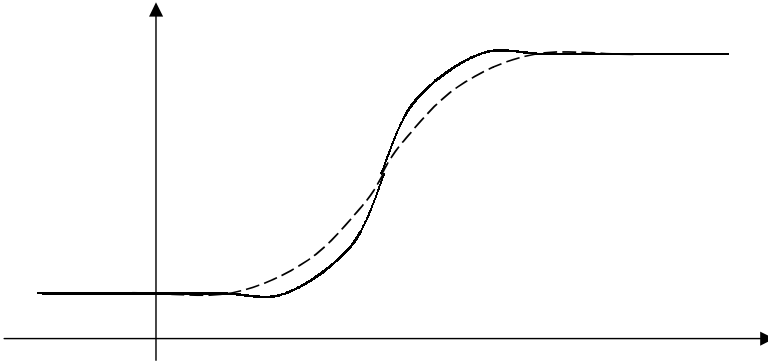

Fig. 1.12. To measure the blur discrimination threshold, first a base blur is chosen (*solid curve*). The ability of the observer to tell apart slightly more blurred edges (*dashed line*) in comparison to this base blur is quantified by psychophysical methods

tested for blur *discrimination* thresholds, i.e. the ability of the observer to tell apart two slightly different amounts of edge blur. As shown in Fig. 1.12, first a base blur (solid curve) is chosen, and the ability of the observer to tell

apart slightly more blurred edges (dashed line) in comparison to this base blur is quantified by psychophysical methods. Figure 1.13 compares model predictions and data from [55] for the effect of exposure duration on blur discrimination thresholds. For both blur perception and discrimination, one
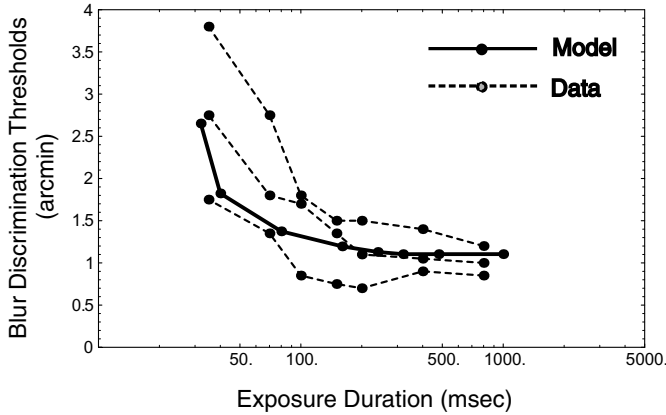

Fig. 1.13. Model predictions (*solid line*) and data (*dashed lines*) of three observers from [55] for blur discrimination threshold as a function of exposure duration. From [44]

observes that an exposure duration on the order of 100 ms is required to reach veridical perception and optimal discrimination of edge blur, and that a good agreement between experimental data and model predictions is found.

Figure 1.14 compares model predictions and data for blur discrimination as a function of base blur. Discrimination thresholds follow a U-shaped function with a minimum value around 1 arcmin. The optics of the eye limits performance for base blurs less than 1 arcmin. For base blurs larger than 1 arcmin, neural factors limit performance.

1.3.4 On and Off Pathways and Edge Localization

Receptive fields of retinal ganglion cells can also be classified as on-center off-surround (Fig. 1.15, left) and off-center on-surround (Fig. 1.15, right). These receptive fields contain two concentric circular regions, called the center and the surround. If a stimulus placed in the center of the receptive field excites the neuron, then a stimulus placed in the surround will inhibit the neuron. Thus the center and the surround of the receptive field have *antagonistic* effects on the neuron. A receptive field whose center is excitatory is called on-center off-surround. Similarly, a receptive field whose center is inhibitory is called off-center on-surround. The outputs of the on-center off-surround cells give rise to the on pathway, and the outputs of the off-center on-surround cells
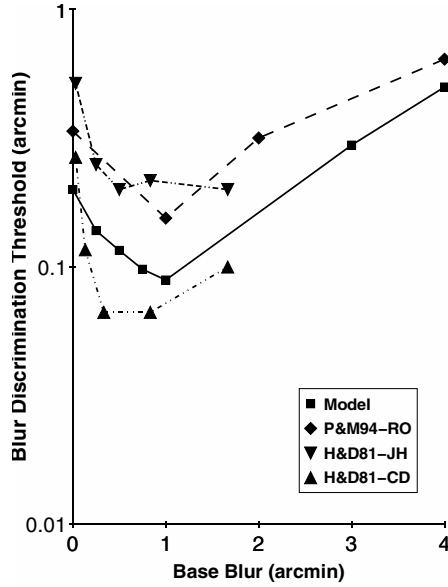
Fig. 1.14. Model predictions and data from [26] (for observers JH and CD) and from [40] (for observer RO) plotting blur discrimination thresholds as a function of base blur. From [44]

give rise to the off pathway. Because the spatial integration of inputs for the P cells is linear, the signals generated by an edge in the on and off pathways will exhibit an odd-symmetry; and their point of balance would correspond to the location of the edge. It has been shown that a contrast-dependent asymmetry exists between the on and off pathways in the human visual system [53]. An implication of this asymmetry is that, if edges are localized based on a comparison of activities in the on and off channels then a systematic mislocalization of the edge should be observed as the contrast of the edge is increased. Indeed, Bex and Edgar [5] showed that the perceived location of an edge shifts towards the darker side of the edge as the contrast is increased. Their data are shown in Fig. 1.16. Negative values on the y-axis indicate that the perceived edge location is shifted towards the darker side of the edge. For a sharp edge (0 arcmin blur), no mislocalization is observed for contrasts ranging from 0.1 to 0.55. However, as the edge blur is increased a systematic shift towards the darker side of the edge is observed. To estimate quantitatively this effect in the model, we introduced an off pathway whose activities consisted of negatively scaled version of the activities in the on pathway. This scaling took into account the aforementioned asymmetry. As a result, as contrast is increased above approximately 0.2, the activities in the off pathway increased slightly more than those in the on pathway. The quantitative predictions of
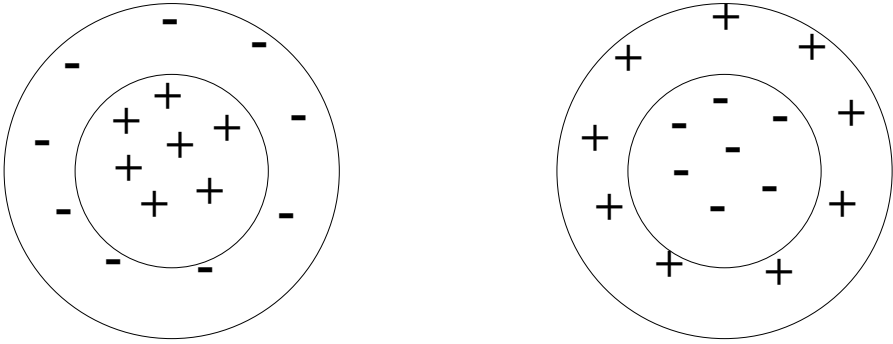
Fig. 1.15. *Left*: On-center off-surround receptive field; *right*: off-center on-surround receptive field. *Plus and minus symbols* indicate excitatory and inhibitory regions of the receptive field, respectively

the model are superimposed on the data in Fig. 1.16. Overall, one can see a good quantitative agreement between the model and the data.

1.3.5 Trade-off Between Spatial and Temporal Deblurring

The aforementioned simulations studied model behavior under the conditions of visual fixation for a static boundary, i.e. when the position of the boundary remains fixed over retinotopic maps. Under these conditions, feedforward retino-cortical signals send blurred boundary information, and gradually post-retinal feedback signals become dominant and construct sharpened representation of boundaries. However, because post-retinal signalling involves positive feedback, at least two major problems need to be taken into consideration:

1) When the positive feedback signals become dominant, the system loses its sensitivity to changes in the input. For example, if the input moves spatially, the signals at the previous location of the input will persist through positive feedback loops and the resulting perception would be highly smeared, similar to pictures of moving objects taken by a camera at long exposure duration. Thus, within a single pathway spatial sharpening comes at the cost of temporal blurring.

2) If left uncontrolled, positive feedback can make the system unstable.

We suggest that the complementary magnocellular pathway solves these problems by rapidly "resetting" the parts of retinotopic map where changes in the input are registered. Accordingly, the real-time operation of the RECOD model unfolds in three phases:

(i) Reset phase: Assume that the post-retinal network has some residual persistent activity due to a previous input. When a new input is applied to
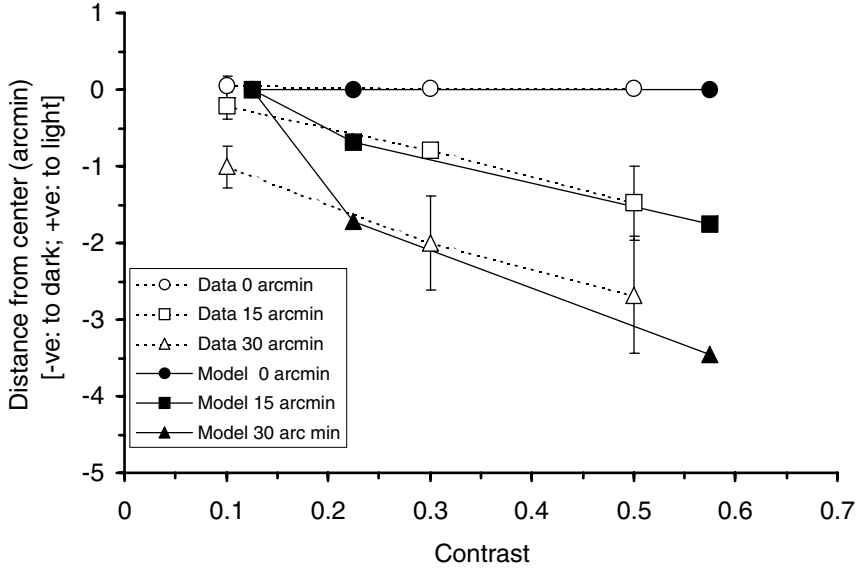
Fig. 1.16. Model predictions and data showing the effect of contrast on the perceived mislocalization of edges with different amounts of blur. The data points are digitized from [5] and represent the mean and the standard error of the mean computed from two observers. From [44]

the RECOD model, the fast-transient neurons respond first. This transient activity inhibits the post-retinal network and removes the persisting residual activity.

(ii) Feedforward dominant phase: The slow-sustained neurons respond next to the applied input and drive the post-retinal network with excitatory inputs.

(iii) Feedback dominant phase: When the activity of the sustained neurons decays from their peak to a plateau, the feedback becomes dominant compared to the sustained feedforward input. This results in the sharpening of the input spatial pattern. Thus, the feedforward reset mode achieves temporal deblurring, and the feedback mode achieves spatial deblurring.

According to the three-phase operation of the model, a single continuous presentation of a blurred edge is necessary for the feedback to sufficiently sharpen the neural image across the retinotopic map. Multiple short exposures cannot achieve the same amount of sharpening as a single long exposure since the post-retinal feedback is reset by the retinal transients. Westheimer [55] measured blur discrimination thresholds for an edge whose blur was temporally modulated in different ways. The reference stimulus was a sharp edge. In the first experiment, the test stimulus was a blurred edge presented alone for durations of 30 ms and 130 ms. Next, the test stimulus was presented as a combination of (i) a sharp edge for 100 ms and a blurred edge for the

next 30 ms, (ii) a blurred edge for the first 30 ms and a sharp edge for the next 30 ms, and (iii) a blurred edge for 100 ms and a sharp edge for the next 100 ms. As shown in Table 1, the RECOD model predicts lower differences in the luminance gradients between the test and reference stimuli for conditions (i) and (ii) above than for a 30 ms presentation of a blurred edge. This gives higher blur discrimination thresholds. Similarly, condition (iii) above yields a lower difference in the luminance gradients between the test and reference stimuli than when the test stimuli is a blurred edge presented for 130 ms.

Table 1.1. Model and data from Westheimer [55] for blur discrimination thresholds (arcmin) obtained with hybrid presentations

	30 ms	130 ms	(i)	(ii)	(iii)
Data	3.8	1.43	7.17	8.56	2.06
Model	2.6	1.2	5.33	5.33	1.44

1.3.6 Perceived Blur for Moving Stimuli

Another way to test the proposed reset phase is to compare model predictions with data on the perception of blur for *moving* stimuli. In normal viewing conditions, moving objects do not appear blurred. Psychophysical studies showed that perceived blur for moving objects depends critically on the exposure duration of stimuli. For example, moving targets appear less blurred than predicted from the visual persistence of static targets when the exposure duration is longer than about 40 ms [10, 28]. This reduction of perceived blur for moving targets was named "motion deblurring" [10].

Model predictions for motion deblurring were tested using a "two-dot paradigm", where the stimulus consisted of two horizontally separated dots moving in the horizontal direction, as shown in the top panel of Fig. 1.17. The middle panel of the figure shows a space-time diagram of the dots' trajectories. The afferent short-latency-transient and long-latency-sustained signals are depicted in the bottom panel of Fig. 1.17 by dashed lines and the gray region, respectively. The sustained activity corresponding to both dots are highly spread over space. However, at the post-retinal level, the interaction between the transient activity generated by the trailing dot and the sustained activity generated by the leading dot results in a substantial decrease of the spatial spread of the activity generated by the leading dot. From Fig. 1.17, one can see that the exposure duration needs to be long enough for the transient activity conveyed by the magnocellular pathway for the trailing dot to spatiotemporally overlap with the sustained activity conveyed by the parvocellular pathway for the leading dot.
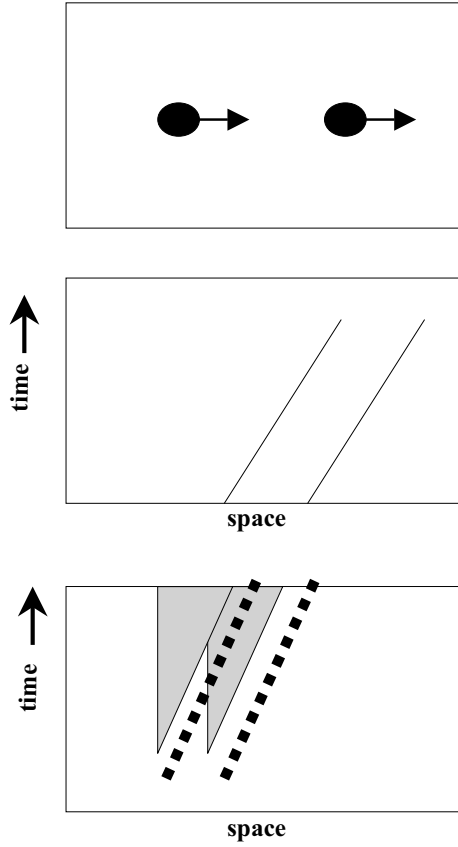
Fig. 1.17. *Top*: Two-dimensional representation of the input. *Arrows* indicate motion. *Middle*: spatiotemporal representation of the input. *Bottom*: superimposed afferent transient and sustained signals

In order to compare model predictions quantitatively with data, Fig. 1.18 plots the duration of perceived blur (calculated as the ratio of the length of perceived blur to the speed) for the leading and the trailing dot, respectively, for two dot-to-dot separations along with the corresponding experimental data [14].

In all cases, when the exposure duration is shorter than 60 msec, no significant reduction of blur is observed and the curves for the leading and trailing dots for both separations largely overlap. The mechanistic explanation of this effect in our model is as follows: due to the relative delay between transient and sustained activities, no spatial overlap is produced when the exposure duration is short. When the moving dots are exposed for a longer duration,
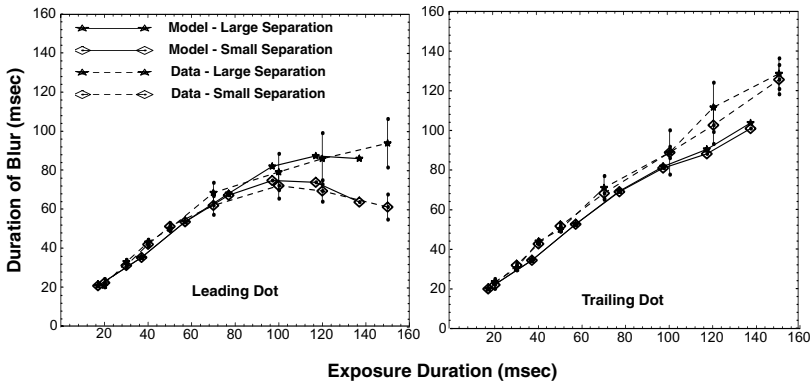
Fig. 1.18. Duration of blur as a function of exposure duration for the leading (left) and trailing (right) dots in the two-dot paradigm for two dot-to-dot separations. From [43]

these two activities overlap and the inhibitory effect of the transient activity on the sustained one reduces the persistent activity from the leading dot. A significant reduction of perceived blur is observed for the leading dot when the dot-to-dot distance is small both in the model and in data. When the dot-to-dot separation is larger, the spatiotemporal overlap of transient and sustained activities is reduced, thereby decreasing the effect of deblurring in agreement with data (Fig. 1.18). For the trailing dot, dot-to-dot separation has no effect on post-retinal activities, and no significant reduction in perceived blur is observed. Quantitatively, the model is in very good agreement with data with the exception of some underestimation for long exposure duration in the case of the trailing dot.

1.3.7 Dynamic Viewing as a Succession of Transient Regimes

Under normal viewing conditions, our eyes move from one fixation point to another, remaining at each fixation for a few hundred millisecons. Our studies show that a few hundred milliseconds is the time required to attain an "optimal" encoding of object boundaries (Figs. 1.11, 1.13, and 1.18). Therefore, the timing of eye movements correlates well with the timing of boundary analysis. We also suggest that these frequent changes in gaze help the visual system remain mainly at its transient regime and thus avoid unstable behavior that would otherwise result from extensive positive feedback loops observed in the post-retinal areas. Within our theoretical framework, the visual and the oculomotor system together "reset" the activities in the positive feedback loops by using the inhibitory fast transient signals originating from the magnocellular pathway.

1.3.8 Trade-off Between Reset and Persistence

If the system is reset by exogenous signals, as suggested above, one needs to consider the problem that may arise because of internal noise: internal noise in the M pathway could cause frequent resets of information processing in areas that compute object boundaries and form. In addition, such rapid undesirable reset cycles may also occur because of small involuntary eye movements as well as because of small changes in the inputs. We suggest that the inhibition from the P-driven system on the M-driven system prevents these resets through a competition between the two systems (see Fig. 1.5). In our simulations reported in the previous sections, for simplicity we did not include sustained on transient inhibition, for both the inputs and the neural activities were noise-free. The proposed competition between the M-driven and the P-driven systems can be tested by using stimuli that activate successively in time spatially nonoverlapping but adjacent regions. The perceptual correlates for such stimuli have been studied extensively in the masking literature [3, 6, 8]. If we label the stimulus whose perceptual and/or motor effects are measured as the "target" stimulus and the other stimulus as the "mask" stimulus (Fig. 1.19), then the condition where the mask is presented in time before the target is called paracontrast. The condition where the mask is presented after the target is called metacontrast [3, 6, 8]. Based on a broad range of masking
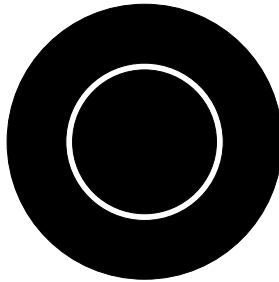

Fig. 1.19. A typical stimulus configuration used in masking experiments. The central disk serves as the target stimulus and the surrounding ring serves as the mask stimulus

data Breitmeyer [7, 6] proposed reciprocal inhibition between sustained and transient channels, and this reciprocal inhibition is also an essential part of the RECOD model. Consider metacontrast: here the aftercoming mask would reset the activity related to the processing of the target. Indeed, a typical metacontrast function is a U-shaped function suggesting that the maximum suppression of target processing occurs when the mask is delayed so that the fast transient activity generated by the mask overlaps in time with the slower sustained activity generated by the target. If the transient activity generated

by the mask can be suppressed by sustained activity, then it should be possible to introduce a second mask (Fig. 1.20) whose sustained activity can suppress the transient activity of the primary mask. This in turn results in the disinhibition of the target stimulus. In support of this prediction, several studies
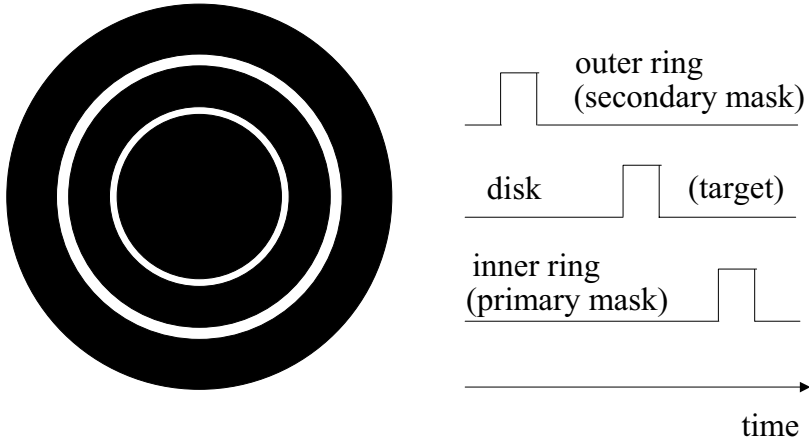

Fig. 1.20. *Left*: modification of the stimulus configuration shown in Fig. 1.19. The second outer ring serves as the secondary mask. *Right*: The temporal order of the stimuli

showed that the second mask allows the recovery of an otherwise suppressed target (e.g. [17]). Furthermore, Breitmeyer et al. [9] showed that the effect of the secondary mask in producing the disinhibition (or recovery) of the target starts when it is presented at about 180 ms prior to the target and gradually increases until it becomes simultaneous with the primary mask. This relatively long range of target recovery provides a time window during which sustained mechanisms can exert their inhibitory influence so as to prevent reset signals generated by noise.

1.3.9 Attention: Real-time Modulation of the Balance Between Reset and Persistence

Having a mechanism to reduce reset signals opens another possibility: modulatory mechanisms can bias the competition in favor of the sustained mechanisms and thereby allow a more persistent and enhanced registration and

analysis of stimuli. We suggest that attention serves that purpose. Although a universally adopted definition of attention does not exist, it is often defined as a selection mechanism whereby resources are focused on certain item(s), location(s), etc. Within the framework of the RECOD model, the reset mechanism curtails cortical activity and therefore attention necessitates a reduction of the reset signals for the attended locations, features, objects, and so on. Similarly, attention can also increase the gain of reset signals for unattended locations and objects. A simple way to achieve this in RECOD is to bias the competition between transient and sustained systems in favor of the sustained system for attended locations, features, and objects; and bias the competition in favor of the transient system for unattended locations, features, and objects. For example, assume that attention primes part of the retinotopic map as illustrated in Fig. 1.21. The model then predicts in agreement with experimental data that attention should increase visible persistence [54], decrease temporal sensitivity [57], increase spatial sensitivity [56], and decrease masking [21, 45, 48]. Similarly, it is predicted that attention should enhance
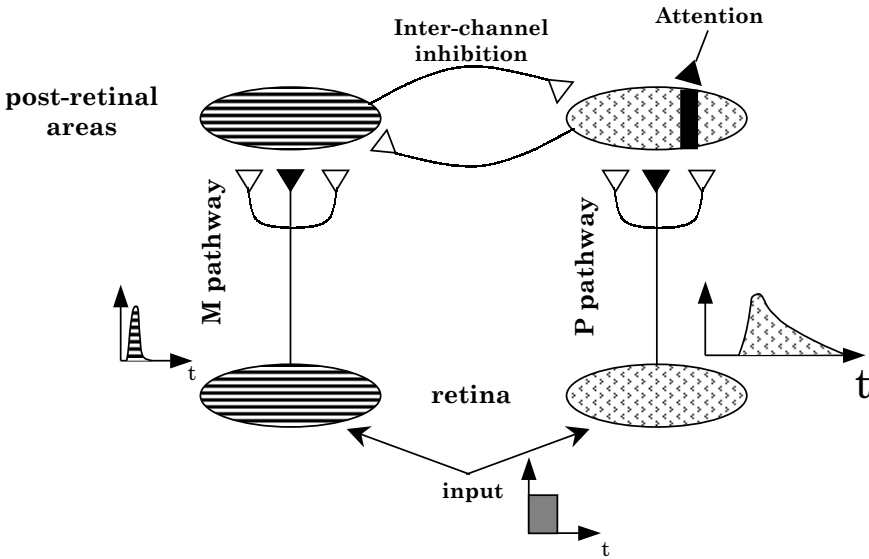

Fig. 1.21. Illustration of attention in RECOD. Priming the activation of the cells in the P pathway biases the competition between sustained and transient systems in favor of the sustained system

target recovery, should increase reaction times to a target in paracontrast, and increase motion blur. These predictions have not been tested.

1.4 Summary

In this chapter we reviewed some fundamental properties of the primate visual system and highlighted maps and pathways as spatiotemporal information encoding and processing strategies. We suggest that maps represent the geometry of the fusion between structure and function in the nervous system, and that the pathways represent complementary aspects of processing whose interactions can solve conflicting requirements arising within a single processing stream. The use of retinotopic and spatial-frequency maps was illustrated by considering the problem of object boundary encoding. The use of parallel, complementary pathways was illustrated by considering how the interactions between magnocellular and parvocellular pathways can resolve the trade-off between spatial and temporal deblurring. We suggested that the interactions between magnocellular and parvocellular pathways play a fundamental role in keeping the system in a succession of transient regimes, thereby avoiding unstable behavior that would result from complex feedback loops that include extensive positive feedback. Finally, we suggested that attention can be viewed as a modulation of the dynamic balance between sustained and transient systems.

Appendix: Fundamental Equations of the Model and Their Neurophysiological Bases

The first type of equation used in the model has the form of a generic Hodgkin–Huxley equation:

$$\frac{dV_m}{dt} = -(E_p + V_m)g_p + (E_d - V_m)g_d - (E_h + V_m)g_h, \qquad (1.1)$$
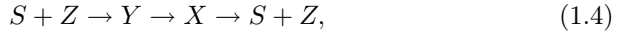
where V_m represents the membrane potential; g_p, g_d, g_h are the conductances for passive, depolarizing, and hyperpolarizing channels, respectively; with E_p, E_d, E_h representing their Nernst potentials. This equation has been used extensively in neural modeling to characterize the dynamics of membrane patches, single cells, as well as networks of cells (rev. [25, 31]). For simplicity, we will assume $E_p = 0$ and use the symbols B, D, and A for E_d, E_h, g_p, respectively, to obtain the generic form for multiplicative or shunting equation (rev. [25]):

$$\frac{dV_m}{dt} = -AV_m + (B - V_m)g_d - (D + V_m)g_h. \qquad (1.2)$$

The depolarizing and hyperpolarizing conductances are used to represent the excitatory and inhibitory inputs, respectively. The second type of equation is a simplified version of Eq. (2), called the additive model, or the leaky-integrator model, where the external inputs influence the activity of the cell not through conductance changes but directly as depolarizing I_d and hyperpolarizing I_h currents yielding the form:

$$\frac{dV_m}{dt} = -AV_m + I_d - I_h. \tag{1.3}$$

Mathematical analyses showed that, with appropriate connectivity patterns, shunting networks can automatically adjust their dynamic range to process small and large inputs (rev. [25]). Accordingly, we use shunting equations when we have interactions among a large number of neurons so that a given neuron can maintain its sensitivity to a small subset of its inputs without running into saturation when a large number of inputs become active. We use the simplified additive equations when the interactions involve few neurons. Finally, a third type of equation is used to express biochemical reactions of the form

$$S + Z \rightarrow Y \rightarrow X \rightarrow S + Z, \tag{1.4}$$

where a biochemical agent S, activated by the input, interacts with a transducing agent Z (e.g. a neurotransmitter) to produce an active complex Y that carries the signal to the next processing stage. This active complex decays to an inactive state X, which in turn dissociates back into S and Z. It can be shown that (see Appendix in Sarikaya et al. [47]), when the active state X decays very fast, the dynamics of this system can be written as:

$$\frac{1}{\tau}\frac{dz}{dt} = \alpha(\beta - z)\gamma Sz, \tag{1.5}$$

with the output given by $y(t) = \frac{\gamma}{\delta}S(t)z(t)$, where s, z, y represent the concentrations of S, Z, and Y, respectively, and γ, δ, α denote rates of complex formation, decay to inactive state, and dissociation, respectively. This equation has been used in a variety of neural models, in particular to represent temporal adaptation, or gain control property, occurring, for example, through synaptic depression (e.g. [1, 13, 22, 24, 37, 38]).

References

1. Abbott L. F., Varela K., Sen K., Nelson S.B. (1997) Synaptic depression and cortical gain control. *Science* **275**:220–223

2. Albright T.D., Desimone R., Gross C.G. (1984) Columnar organization of directionally selective cells in visual area MT of the macaque. *J. Neurophysiol.* **51**:16–31

3. Bachmann T. (1994) *Psychophysiology of Visual Masking: The Fine Structure of Conscious Experience.* Nova Science, New York

4. Baron M., Westheimer, G. (1973) Visual acuity as a function of exposure duration. *J. Opt. Soc. Am.* **63**:212–219

5. Bex P.J., Edgar G.K. (1996) Shifts in perceived location of a blurred edge increase with contrast. *Perception and Psychophysics* **58**:31–33

6. Breitmeyer B.G. (1984) *Visual masking: An Integrative Approach.* Oxford University Press, Oxford

7. Breitmeyer B.G., Ganz, L. (1976) Implications of sustained and transient channels for theories of visual pattern masking, saccadic suppression, and information processing. *Psychological Rev.* **83**:1–36

8. Breitmeyer B.G., Öğmen H. (2000) Recent models and findings in visual backward masking: A comparison, review, and update. *Perception and Psychophysics* **62**:1572–1595

9. Breitmeyer B.G., Rudd M., Dunn K. (1981) Metacontrast investigations of sustained-transient channel inhibitory interactions. *J. of Exp. Psych: Human Perception and Performance* **7**:770–779

10. Burr D. (1980) Motion smear. *Nature* **284**:164–165

11. Burr D.C., Morgan, M.J. (1997) Motion deblurring in human vision. *Proc. R. Soc. Lond. B* **264**:431–436

12. Carandini M., Heeger D.J. (1994) Summation and division by neurons in primate visual cortex. *Science* **264**:1333–1336

13. Carpenter G. A., Grossberg S. (1981) Adaptation and transmitter gating in vertebrate photoreceptors. *J. of Theor. Neurobiology* **1**:1–42

14. Chen S., Bedell H.E., Öğmen H. (1995) A target in real motion appears blurred in the absence of other proximal moving targets. *Vision Res.* **35**:2315–2328

15. Croner L.J., Kaplan E. (1995) Receptive fields of P and M ganglion cells across the primate retina. *Vision Res.* **35**:7–24

16. De Monasterio F.M. (1978) Properties of concentrically organized X and Y ganglion cells of macaque retina. *J. Neurophysiol.* **41**:1394–1417

17. Dember W.N., Purcell D.G. (1967) Recovery of masked visual targets by inhibition of the masking stimulus. *Science* **157**:1335–1336

18. De Valois K.K. (1977) Spatial frequency adaptation can enhance contrast sensitivity. *Vision Res.* **17**:209–215

19. De Valois R.L., De Valois K.K. (1990) *Spatial Vision.* Oxford University Press, New York

20. De Valois K.K., Switkes E. (1980) Spatial frequency specific interaction of dot patterns and gratings. *Proc. Nat. Acad. Sci. USA* **77**:662–665

21. Enns J.T., DiLollo V. (1997) Object substitution: A new form of masking in unattended visual locations. *Psychological Science* **8**:135–139

22. Gaudiano P. (1992) A unified neural network of spatio-temporal processing in X and Y retinal ganglion cells. 2: Temporal adaptation and simulation of experimental data. *Biol. Cybern.* **67**:23–34

23. Georgeson M.A. (1994) From filters to features: location, orientation, contrast and blur. *CIBA Foundation Symposia* **184**:147–169

24. Grossberg S. (1972) A neural theory of punishment and avoidance, II: Quantitative theory. *Mathematical Biosciences* **15**:253–285

25. Grossberg S. (1988) Nonlinear neural networks: Principles, mechanisms and architectures. *Neural Networks* **1**:17–61
26. Hamerly J.R., Dvorak, C.A. (1981) Detection and discrimination of blur in edges and lines. *J. Opt. Soc. Am.* **71**:448–452
27. Hendry S.H.C., Reid, R.C. (2000) The koniocellular pathway in primate vision. *Annu. Rev. Neurosci.* **23**:127–153
28. Hogben J.H., Di Lollo V. (1985) Suppression of visible persistence in apparent motion. *Perception and Psychophysics* **38**:450–460
29. Hood D. (1973) The effects of edge sharpness and exposure duration on detection threshold. *Vision Res.* **13**:759–766
30. Hubel D.H., Wiesel T.N. (1968) Receptive fields and functional architecture of monkey striate cortex. *J. Physiol. London* **195**:215–243
31. Koch C., Segev I. (1989) *Methods in Neuronal Modeling.* MIT Press, Cambridge, MA
32. Livingstone M., Hubel, D. (1988) Segregation of form, color, movement, and depth: Anatomy, physiology, and perception. *Science* **240**:740–749
33. Maunsell J.H.R., Gibson J.R. (1992) Visual response latencies in striate cortex of the macaque monkey. *J. Neurophysiol.* **68**:1332–1344
34. Milner A.D., Goodale M.A. (1995) *The Visual Brain in Action.* Oxford University Press, Oxford
35. Mishkin M., Ungerleider L.G., Macko, K.A. (1983) Object vision and spatial vision: Two cortical pathways. *Trends in Neurosciences* **6**:414–417
36. Nowak L.G., Munk M.H.J., Girard P., Bullier J. (1995) Visual latencies in areas V1 and V2 of the macaque monkey. *Visual Neuroscience* **12**:371–384
37. Öğmen H. (1993) A neural theory of retino-cortical dynamics. *Neural Networks* **6**:245–273
38. Öğmen H., Gagné S. (1990) Neural models for sustained and on-off units of insect lamina. *Biol. Cybern.* **63**:51–60
39. Öğmen H., Breitmeyer B.G., Melvin R. (2003) The what and where in visual masking. *Vision Res.* **43**:1337–1350
40. Pääkkönen A.K., Morgan M.J. (1994) Effect of motion on blur discrimination. *J. Opt. Soc. Am. A* **11**:992–1002
41. Pizlo Z. (2001) Perception viewed as an inverse problem. *Vision Res.* **41**:3145–3161
42. Purpura K., Tranchina D., Kaplan E., Shapley R.M. (1990) Light adaptation in primate retina: Analysis of changes in gain and dynamics of monkey retinal ganglion cells. *Visual Neuroscience* **4**:75–93
43. Purushothaman G., Öğmen H., Chen S., Bedell H.E. (1998) Motion deblurring in a neural network model of retino-cortical dynamics. *Vision Res.* **38**:1827–1842
44. Purushothaman G., Lacassagne D., Bedell H.E., Öğmen H. (2002) Effect of exposure duration, contrast, and base blur on coding and discrimination of edges. *Spatial Vision* **15**:341–376
45. Ramachandran V.S., Cobb S. (1995) Visual attention modulates metacontrast masking. *Nature* **373**:66–68
46. Salin P.-A., and Bullier J. (1995) Corticocortical connections in the visual system: structure and function. *Physiological Reviews* **75**:107–154
47. Sarikaya M., Wang W., Öğmen H. (1998) Neural network model of on-off units in the fly visual system: simulations of dynamic behavior. *Biol. Cybern.* **78**:399–412

48. Shelley-Tremblay J., Mack A. (1999) Metacontrast masking and attention. *Psychological Science* **10**:508–515
49. Stecher S., Sigel C., Lange R.V. (1973) Composite adaptation and spatial frequency interactions. *Vision Res.* **13**:2527–2531
50. Tolhurst D.J. (1972) Adaptation to square-wave gratings: Inhibition between spatial frequency channels in the human visual system. *J. Physiol.* **226**:231–248
51. Tootell R.B.H., Silverman M.S., De Valois R.L. (1981) Spatial frequency columns in primary visual cortex. *Science* **214**:813–815
52. Tootell R.B.H., Silverman M.S., Hamilton S.L., De Valois R.L., Switkes E. (1988) Functional anatomy of macaque striate visual cortex. 3. Color. *J. Neurosci.* **8**:1569–1593
53. Virsu V., Laurinen, P. (1977) Long-lasting afterimages caused by neural adaptation. *Vision Res.* **17**:853–860
54. Visser T.A., Enns J.T. (2001) The role of attention in temporal integration. *Perception* **30**:135–145
55. Westheimer G. (1991) Sharpness discrimination for foveal targets. *J. Opt. Soc. Am.* **8**:681–685
56. Yeshurun Y., Carrasco M. (1998) Attention improves or impairs visual performance by enhancing spatial resolution. *Nature* **396**:72–75
57. Yeshurun Y., Levy L. (2003) Transient spatial attention degrades temporal resolution. *Psychological Science* **14**:225–231
58. Youping X., Yi W., Felleman D.J. (2003) A spatially organized representation of colour in macaque cortical area V2. *Nature* **421**:535–539
59. Zeki S. (1997) The color and motion systems as guides to conscious visual perception. *Cerebral Cortex* **12**:777–809