Chapter 7
Theoretical Models of the Motion Aftereffect

George Mather and John Harris

In a rather short chapter, Holland (1965) reviewed a number of early theoretical explanations of the motion aftereffect (MAE) from the 1800s up to the 1950s, such as eye movement and blood flow theories, but he did not discuss the new theory that emerged in the early 1960s inspired by cortical cell physiology. This theory placed the MAE in a unique position among perceptual phenomena in terms of the directness of the proposed links between cortical cell activity and perception. It is founded on two psychophysical linking hypotheses: (1) that perception of motion is mediated by some form of comparison between the responses of cells in the visual system sensitive to different directions; and (2) that following adaptation to a moving stimulus there is a change in the responsiveness of these cells, so that cells tuned to motion directions congruent with the adapting stimulus show a reduction in response relative to cells tuned to other directions.

In the spirit of Brindley (1970), we should first confirm the plausibility of these hypotheses by correlating properties of neural events with corresponding properties of perceptual phenomena. This is a straightforward task in the case of motion perception and aftereffects. Data from direct cell recordings show that the middle temporal area (MT) in primates is particularly rich in motion-sensitive cells (see chapter 6). Human brain imaging studies show high activity in a corresponding region of cortex—the occipitotemporal parietal junction—in the presence of MAEs (Tootell et al., 1995), and closed head injuries in the same area of cortex lead to impaired motion perception (Vaina et al., 1990). It is also not difficult to find more specific evidence supporting the involvement of cortical cells in the MAE. For example, classic MAEs are confined to the area of retina exposed to the adapting stimulus, a property that can be related to the restricted receptive field of cortical cells; the perceptual phenomenon is also usually short-lived, in agreement with data on changes in neural responsiveness following adaptation; and the degree of interocular transfer of the aftereffect can be related to the binocularity of cortical cells, and may be used to infer the probable sites of adaptation, as discussed below

and in chapter 4. There are exceptions to these clear psychophysical links which are theoretically significant (e.g., von Grünau, 1986; Masland, 1969), but on the basis of a large body of evidence, much of it surveyed in this book, we can accept the two hypotheses as a firm basis for constructing theoretical models. The first section begins by discussing the first, and simplest modern theoretical model of MAE, and then discusses a more complex model that is also able to accommodate other phenomena in motion perception. Section 7.2 examines the functional significance of perceptual adaptation in relation to MAEs.

7.1 Models of Direction Coding

The MAE has both a direction and an apparent speed. However, theoretical models have restricted themselves to explaining the directional properties of the effect. Apparent speed has been used predominantly as a measure of MAE magnitude, since it correlates very well with duration (Pantle, 1974). As there have been no attempts to built explicit assumptions about velocity coding into explanatory models of the MAE, this section deals only with models of direction coding.

7.1.1 Opponent Process Coding

Precisely how do the two hypotheses above permit an explanation of the MAE? Sutherland (1961) proposed the first minimally sufficient model of the MAE—the ratio or "opponent-process" model:

> Hubel and Wiesel (1959) have, however, found cells which respond differentially according to the direction in which a stimulus is moved across the retina. If direction of movement is coded in single cells in human beings, adaptation in these cells might clearly underly [*sic*] the after-effect of movement. Once again the direction in which something is seen to move might depend upon the ratios of firing in cells sensitive to movement in different directions, and after prolonged movement in one direction a stationary image would produce less firing in the cells which had just been stimulated than normally [*sic*], hence apparent movement in the opposite direction would be seen to occur. (p. 227)

Sutherland's prediction of adaptation effects in single visual cells was first confirmed by Barlow and Hill (1963), who measured responses in rabbit retinal ganglion cells, and later confirmed by a number of workers recording from cat and monkey cortical cells (see chapter 6). Barlow and Hill (1963) themselves concluded that "the after-effects of motion may result from the temporary imbalance of the maintained discharges of cells responsive to opposite directions" (p. 1346).

There is a subtle difference in wording between Sutherland's and Barlow and Hill's proposals, in that the former deals with comparisons between cells tuned to "different" directions, and the latter deals with comparisons between cells tuned to "opposite" directions. The opponent-process account has become the standard explanation of the MAE. Direction-selective cells tuned to opposite directions provide paired inputs to a comparator cell, one excitatory and the other inhibitory. Perceived direction is said to depend on the difference between the outputs of the oppositely tuned detectors, signaled by the comparators. The sign of the difference in detector output is crucial, of course, since this specifies direction sense. For example, assume that detectors tuned to upward motion provide excitation at the comparator, while detectors tuned to downward motion provide inhibition. Net excitation at the comparator then signifies upward motion, and net inhibition signifies downward motion. However, it is not feasible physiologically for a comparator cell to signal both excitation and inhibition (i.e., signed differences) over a wide dynamic range. The solution to this kind of problem, as we know from studies of retinal ganglion cells that signal intensity differences, is to have separate comparator cells supply the positive and negative portions of the difference signal as positive responses. Some comparators supply the positive half of the response (i.e., are excited by upward motion and inhibited by downward motion) and others provide the negative half of the response (i.e., are inhibited by upward motion and excited by downward motion). This scheme is illustrated in figure 7.1.

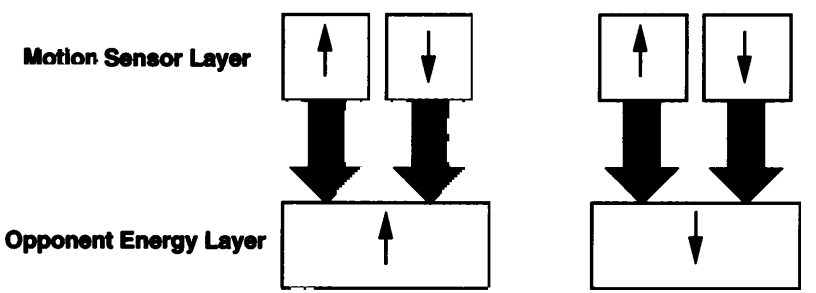

Figure 7.1
Simple opponent-process model of direction coding. Direction-selective motion sensors (upper layer) provide paired inputs to opponent-energy units (lower layer). One input is excitatory (light gray), and the other input is inhibitory (dark gray). On the left, sensors tuned to upward motion provide excitation and sensors tuned to downward motion provide inhibition, so the opponent-energy unit produces a positive response to upward motion. On the right, excitatory and inhibitory inputs are reversed, so that the opponent-energy unit produces a positive response to downward motion.

The model contains two layers of units. Motion sensors in the first layer provide initial measurements of motion energy. Their opposed outputs feed a pair of units in the opponent-energy layer, one of which provides a positive signal to upward motion, while the other provides a positive signal to downward motion. Motion in one direction is perceived when the output of the opponent-energy unit signaling that direction exceeds some internal threshold. The opponent-energy units correspond to those proposed by Adelson and Bergen's (1985), partly on the basis that "adaptation phenomena such as the MAE suggest that motion perception involves the balance between opposing leftward- and rightward-motion signals" (p. 293). Note that responses at the sensor layer of this scheme interact competitively, but responses at the opponent-energy layer do not.

In principle, adaptation could arise at the sensor layer, or at the opponent-energy layer, or at both. What are the predicted effects of adaptation in the two layers? We assume that adaptation in either layer has two consequences for cell activity. First, the resting level of the affected cell is depressed. Second, the amount of stimulation required to reach a particular level of response in the cell is elevated. Figure 7.2 illustrates the pattern of responses in the two layers during an MAE experiment. The upper row of graphs plots the output of units in the sensor layer sensitive to upward and downward motion, and the lower row of graphs plots the output of units in the opponent-energy layer. All responses are shown relative to a resting level of activity (which could also represent the small response to a nondirectional stimulus). As indicated, the response of each unit in the opponent-energy layer is given by the sum of its resting level and the difference between the responses of two sensor units. Responses in this layer that exceed some minimum magnitude, shown by the dashed line at threshold, lead to the perception of motion.

Before adaptation and in the absence of motion (figure 7.2a), the system is in equilibrium, with all units at resting level. During adaptation to upward motion (figure 7.2b), the upward sensor UP_s responds strongly, but the downward sensor $DOWN_s$ remains at resting level. This leads to an above-threshold response from the upward opponent-energy unit UP_o and a suppressed response from the downward opponent-energy unit DOWNo. Consider first the result of adaptation that is confined only to the *sensor* layer (figure 7.2c). The resting level of the upward sensor will be depressed, whereas the resting level of the downward sensor will be unaffected. This difference will be reflected in the outputs of the opponent-energy units, with the upward unit showing a depressed response and the downward unit showing an above-threshold response that should lead to perception of an MAE. Now consider the consequences of adaptation that is confined only to the *opponent-energy* layer (figure 7.2d).

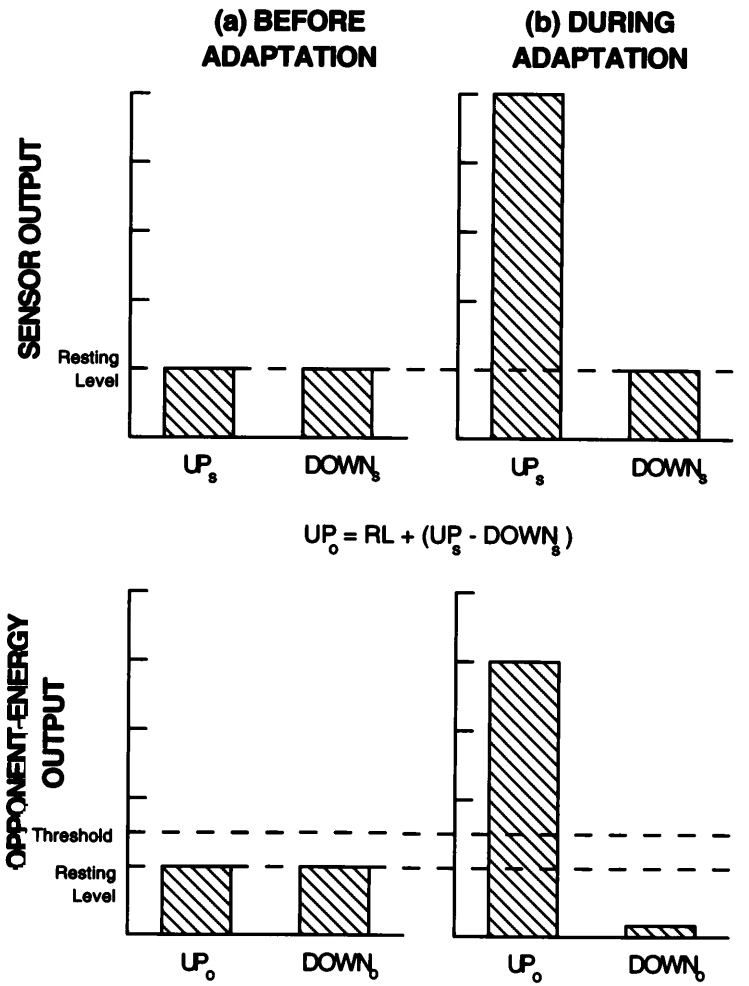

Figure 7.2

Explanation of the motion aftereffect (MAE), according to the opponent-process model in figure 7.1. The upper row of graphs shows the output of motion sensors tuned to upward motion (UP_s) and to downward motion ($DOWN_s$) in different stimulus conditions. Sensor output in the absence of motion is shown by the broken line. Each graph in the lower row shows the output of opponent-energy units connected to the sensors depicted in the graph immediately above. Opponent-energy units signaling upward motion (UP_o) receive positive inputs from upward sensors and negative inputs from downward sensors, and vice versa for opponent-energy units signaling downward motion ($DOWN_o$). It is assumed that each opponent-energy unit has a resting level of response, and that motion is seen only when opponent-energy output exceeds a threshold value (broken lines). Different columns show sensor and opponent-energy output (a) before adaptation and in the absence of motion; (b) during adaptation to upward motion; (c) after adaptation and in the absence of motion,

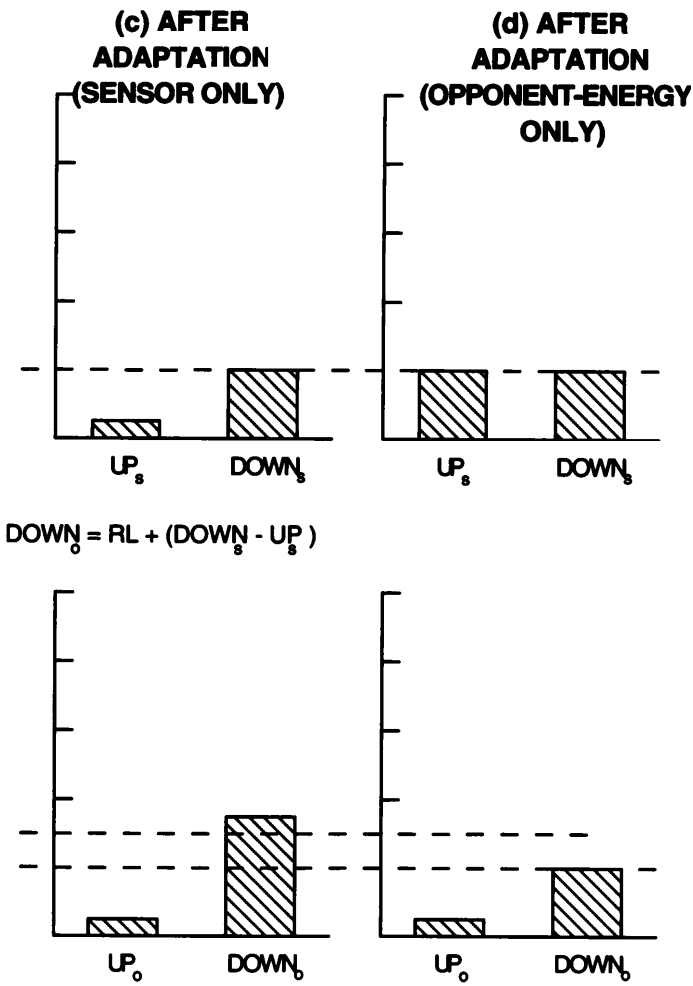

Figure 7.2 (continued)
assuming that only sensor output is depressed; (d) after adaptation and in the absence of motion, assuming that only opponent-energy output is depressed. The graphs reveal that only sensor adaptation can lead to above-threshold opponent-energy responses in the absence of motion, that is, MAEs. See text for more details.

Following adaptation there will be no imbalance in sensor outputs, but despite this the opponent-energy unit that was active during adaptation will show a suppressed response, whereas the unit that was not active will be unaffected. Neither opponent-energy unit will respond above threshold, so there will be no MAE, but the depressed output in the adapted unit should lead to a loss of sensitivity to the adapted direction, since more stimulation will be required to exceed threshold than before adaptation, perhaps reflected in higher motion detection thresholds. In reality, adaptation may be present in both layers, but the main point is that only adaptation in the sensor layer is associated with an MAE.

The very existence of the MAE points to the presence of adaptation in the sensor layer. Is there any evidence for the presence of adaptation in the opponent-energy layer? Raymond (1993a) measured motion coherence thresholds for motion in the four cardinal directions (up, down, left, right) following adaptation to rightward motion. She found significantly reduced sensitivity to rightward motion, but no significant changes in sensitivity to the other three directions. This result can be explained by the opponent-process model if we assume that the obtained coherence threshold elevation mainly reflected adaptation at the opponent-energy layer in figures 7.1 and 7.2. In a second experiment, Raymond (1993b) found that coherence thresholds for unidirectional motion were raised by approximately 22 percent after bidirectional adaptation, but were raised by 47 percent after unidirectional adaptation. According to the model, unidirectional adaptation should drive units in both layers strongly, whereas bidirectional adaptation should drive only sensors (the opposite sensor signals tend to cancel out at opponent-energy units). The obtained difference between unidirectional and bidirectional adaptation effects may therefore reflect adaptation in opponent-energy units. The effect reported by Raymond and Braddick (1996; see chapter 5, figure 5.6) can also be explained by adaptation at the opponent-energy layer.

7.1.2 Two-Dimensional Models

Despite its success in accounting for some basic properties of the MAE and motion adaptation, the opponent-process model sketched above has serious limitations. First, it cannot accommodate the high-level MAE phenomena already described in chapter 5 (e.g., second-order motion, multivectorial MAEs). Second, it is inherently one-dimensional, since it codes direction only along a single axis, but there are strong grounds, both empirical and theoretical, for believing that human motion perception involves two-dimensional (2-D) analysis (e.g., Adelson and Movshon, 1982). The addition of a third layer to the model permits 2-D interactions between motion signals which can potentially overcome many of these limitations. The model proposed by Wilson et al. (1992)

and Wilson and Kim (1994) contains three layers of units, the first two of which correspond to the sensor and opponent-energy layers sketched in figure 7.1. The third layer contains integrator units which receive both excitatory and inhibitory inputs from opponent-energy units tuned to a wide range of directions, in order to compute global motion direction. Figure 7.3 is a simple illustration of the model. The top row depicts the preferred direction of units in the sensor layer, the middle row depicts units in the opponent-energy layer, and the bottom row depicts units in the integrator layer. Units tuned to directions within ± 120 degrees from vertical are shown, with each unit having directional tuning of ± 11 degrees.

Thus units in the first two layers correspond to the units depicted in figure 7.1. Each integrator unit in the third layer sums inputs from a range of opponent-energy units signaling directions within a range of ± 120 degrees. For illustration, only connections to the integrator tuned to upward motion are shown. Opponent-energy inputs within ± 75 degrees are excitatory (light-gray connections), and the remainder are inhibitory (black connections), weighted so that the maximum response in the integrator unit layer will be from a unit tuned to the vector sum direction of the input activity. There are recurrent inhibitory (feedback) connections between integrator units, so that each integrator unit inhibits other units with preferred directions differing by between ± 45 degrees and ± 120 degrees. The figure illustrates the inhibitory connections feeding back from the upward integrator (black connections). This inhibition generates a form of "winner-take-all" interaction, and the restriction of interactions to ± 120 degrees allows for more than one winner to be computed, that is, motion transparency.

Wilson and Kim (1994) proposed that the opponent-energy layer contains both "first-order" and "second-order" units. In first-order stimuli (Cavanagh and Mather, 1989; Chubb and Sperling, 1988) the motion signal is carried by stable differences in intensity (e.g., drifting luminance gratings). In second-order stimuli there are no stable intensity differences correlated with the motion signal. Motion is carried by differences in texture properties (e.g., contrast, spatial scale, temporal modulation).[1] Wilson and Kim tentatively identified the opponent-energy layer of the model with cells in cortical areas V1 and V2, and identified the integrator layer with cells in cortical area MT.

7.1.3 Multiple Sites of Adaptation

Since this model is built from the same sensor and opponent-energy units as those in figure 7.1, it can provide the same explanation for simple MAEs, if we assume that adaptation occurs in the sensor layer (H. R. Wilson, personal communication, 1997). Selective adaptation in this layer

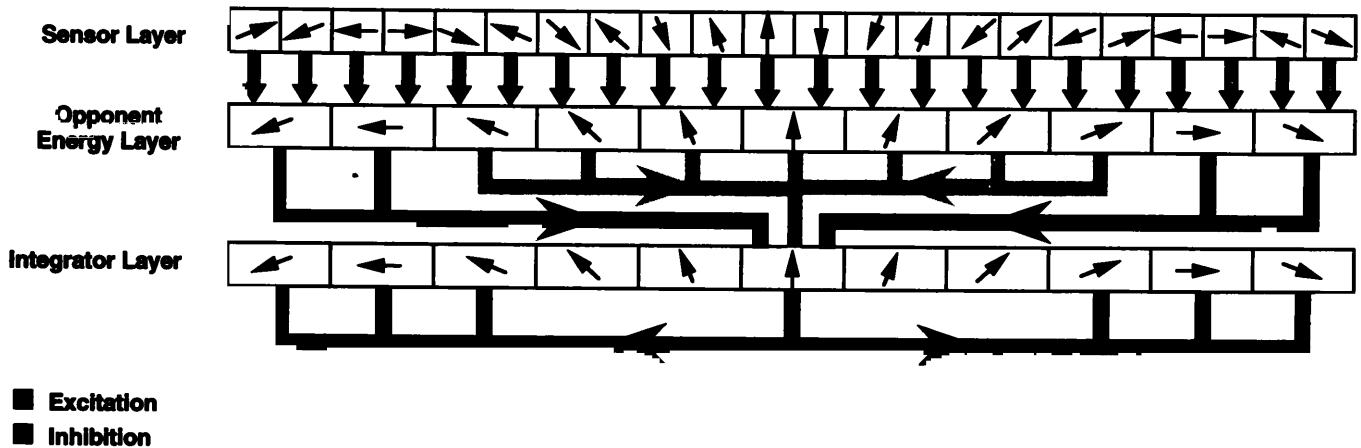

Figure 7.3
Three-layer model of motion processing proposed by Wilson et al. (1992) and Wilson and Kim (1994). The first two layers of the model (sensor and opponent-energy) contain arrays of units that correspond to those described in figure 7.1. Each unit is tuned to a range of directions spanning ± 11 degrees. The third layer contains an array of integrator units, each receiving inputs from a range of opponent-energy units spanning directions within ± 120 degrees. Opponent-energy inputs within ± 75 degrees are excitatory (light-gray connections), and those in the range of 75 to 120 degrees are inhibitory (black connections). In addition, there are recurrent inhibitory connections between integrators, so that each integrator inhibits others tuned to a range of directions spanning ± 5 to ± 120 degrees. Note that the full array of units in each layer spans all directions around the clockface. For the purposes of illustration the figure only shows those units that provide ascending inputs to the integrator unit tuned to upward motion. This unit also receives recurrent inhibition from other integrators, but these too are not shown.

will be expressed as an imbalance between excitation and inhibition in the opponent-energy layer, leading to an MAE. The presence of integrators in the three-layer model introduces two more potential sources of MAE signals. Recall that in the original opponent-process model selective adaptation in the opponent-energy layer alone could not lead to an MAE. In the three-layer model, selective adaptation in the opponent-energy layer could potentially result in an MAE. For example, upward adaptation would depress the response of the upward opponent-energy unit, but would leave the downward unit unaffected (figure 7.2d). The resulting imbalance between excitation and inhibition arriving from the opponent-energy layer may be sufficient to generate a motion signal at the integrators. Indeed, as far as integrators are concerned, it should not matter whether the imbalance arises from sensor adaptation or from opponent-energy adaptation. MAEs could also arise from adaptation that was confined only to the integrators. In this case the resultant change in the pattern of recurrent inhibition between integrators may be sufficient to generate a motion signal.

Initial results of computational modeling (H. R. Wilson, personal communication) indicate that adaptation of integrator units can certainly explain changes in perceived plaid coherence, and changes in the perceived direction of moving stimuli. There is some psychophysical evidence that integrator adaptation also contributes to MAEs. Verstraten et al. (1994a, p. 356) measured MAEs following adaptation to two transparent motion fields which individually generated MAEs of different duration. They reasoned that if the resulting MAE arose from adaptation of individual responses to each field, then the MAE should change direction as the effect of the weaker adapting component disappeared. No change in direction was reported, so the aftereffect must have been generated at a site after the individual responses had been combined. Van Wezel, Verstraten, et al. (1994b) measured motion discrimination thresholds and MAE durations using a checkerboard pattern in which alternating checkers contained texture drifting in opposite directions. The two measures were differentially affected by checker size, leading the authors to conclude that the adaptation effect occurred at an integration stage which covers a much greater retinal area than that occupied by the receptive fields of individual sensors.

Recall from chapter 5 that consistent differences have been reported between static MAEs and flicker MAEs, leading a number of workers to conclude that they reflect adaptation at different levels of motion analysis (see chapter 5, table 5.1). To take one example, Nishida et al. (1994) measured the relative duration of monocular and interocular MAEs using static and flickering tests. For static tests, interocular MAEs lasted only 30 to 50 percent as long as monocular MAEs, but there was little difference between monocular and interocular conditions for flicker MAEs. Nishida

et al. argued that more complete interocular transfer indicates adaptation at higher levels of processing, and on this basis argued that static MAEs reflect adaptation at low-level detectors and flicker MAEs reflect adaptation at high-level integration. They speculatively identified the latter with cells in cortical area MT. A simple application of the authors' interpretation to the three-layer model would identify static MAEs with adaptation at the sensor layer, and flicker MAEs with adaptation at the integrator layer.

The empirical differences between static and flicker MAEs shown in table 5.1 are certainly consistent with this idea. For example, integrator units are likely to have larger receptive fields, show less spatiotemporal specificity, and be more binocular than sensor units. However, this interpretation does beg the following question. Why should static test stimuli favor the contribution of sensor adaptation to the visible MAE, and flickering test stimuli favor the contribution of integrator adaptation? The argument that sensors are sensitive to stationary test patterns and integrators are sensitive only to dynamic patterns (cf. McCarthy, 1993; Nishida and Sato, 1995) is not tenable. Any sensor response that leads to a motion percept must necessarily generate a directional signal in integrators, so anything that sensors "see," integrators must "see" also. It is fair to assume that dynamic test stimuli will drive motion sensors tuned to many directions much more effectively than will static tests. Perhaps differences between the responses of adapted and unadapted units are greatest at relatively low response levels (cf. response normalization in spatial vision), so dynamic tests minimize the contribution of sensor adaptation, and static tests maximize its contribution. There is no clear answer to this question at present, so further research is needed.

We have seen that some MAE phenomena can be attributed to sensor adaptation, and others can be attributed to integrator adaptation. If this is the case, the middle layer of opponent-process units in the model may be superfluous—we could omit the top layer from figure 7.3 and relabel the opponent-energy layer as the sensor layer. Adaptation-induced differences in sensor output would result in imbalances between inhibition and excitation arriving directly at the integrators. However, it is not possible to determine the significance of the opponent-energy layer without detailed computational analysis of the model. In the meantime, the issue remains open. Adaptation-induced changes in global motion thresholds were earlier (section 7.1.1) attributed to the opponent-energy layer, and it is not clear how well the model can account for such effects without this layer.

With these caveats in mind, the general motion-processing scheme outlined in figure 7.4 includes only the sensor and integrator layers of the model in figure 7.3, since these seem the minimum necessary to accommodate much of the MAE data. A few points are worthy of emphasis.

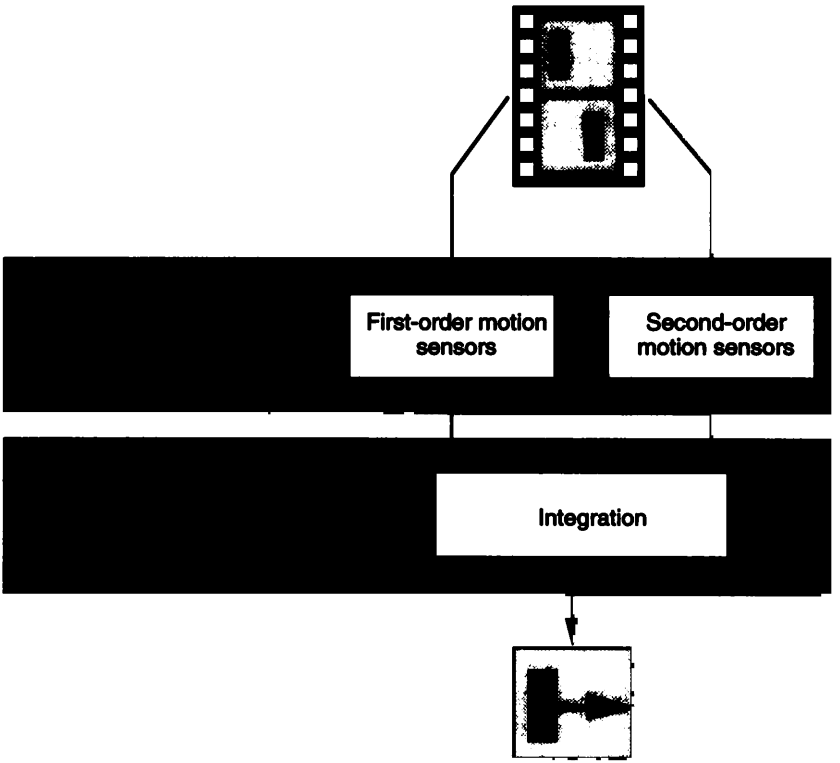

Figure 7.4
Hypothetical architecture for motion processing, incorporating motion aftereffect (MAE) phenomena. Only two of the three layers from the model in figure 7.3 are shown, since these seem the minimum necessary to accommodate the data. Properties associated with adaptation in the two layers are shown inside the boxes.

First, the perceptual manifestation of an MAE must represent the combined effect of adaptation at both sites. Neurons in the two layers are likely to differ in their response properties, such as receptive field size and binocularity (Wilson and his colleagues have speculated that sensors are located in V1, and integrators in MT), so their relative contribution to the resultant MAE should depend on stimulus conditions (i.e., the nature of adapting and test stimuli), leading to the different properties listed in the figure. Second, in the figure, static MAEs are attributed to sensors, and dynamic MAEs are attributed to integrators, but as we have seen there is as yet no coherent account for this division. Third, the difference in duration between sensor and integrator adaptation reflects the conclusion in chapter 3, section 3.1, that there is an association between binocularity, spatial specificity, and duration of MAEs. Fourth, given the empirical

properties of adaptation effects from second-order motion (e.g., similarities with dynamic MAE properties; see chapter 5), they may predominantly reflect integrator adaptation.

7.1.4 Summary

The original opponent-process model attributed the MAE to adaptation at a single neural site (motion sensors). This model can no longer offer an adequate account of the phenomenon. Instead, the complex pattern of MAE data implicates a motion-processing system that involves at least two stages of analysis, incorporating 2-D interactions, with the potential for adaptation at both stages. Models of motion perception containing either two or three layers of analysis appear capable of accommodating much of the MAE data, if assumptions about adaptation are included. However, a firm conclusion on the most capable model must await detailed computational research. Important questions remain regarding the explanation of differences between static and dynamic MAEs.

7.2 What is the Functional Role of Selective Adaptation?

As the rest of this book shows, selective adaptation has been one of the most important tools in the study of the early stages of processing in human vision. It has been critical not only in helping to identify the range of stimulus attributes which are independently processed and to characterize the tuning of the underlying mechanisms but also in providing strong links between psychophysic studies on people and physiologic studies on animals. However, despite its undoubted utility in experiments, the phenomenon of adaptation itself remains puzzling. Why should it occur at all?

The reason for the puzzle is as follows. One might suppose that it is important for vision, or any other sensory system, to provide its owner with as veridical a picture of the world as possible. Consequently, any design fault which introduced distortions would be maladaptive and be weeded out by natural selection. However, aftereffects seem to break this rule: they show in a very immediate fashion that the visual system can produce profoundly distorted messages about the world, so that an object that is physically upright may appear to tilt and one that is physically stationary may appear to be moving. On the face of it, this appears to be an example of bad design which should have been removed during evolution.

Nevertheless, the idea that aftereffects result from the "fatigue" or "satiation" of visual neurons has been a pervasive one (Kohler and Wallach, 1944), driven perhaps by an implicit analogy between the supposed effects of continued stimulation on sensory nerves and those of continued

exercise on skeletal muscles. Fatigue might reflect, for example, the inability of neurons to continue to produce neurotransmitter at high concentrations for long periods. However, there are several lines of evidence against this idea. First, visual aftereffects can result from very short adapting exposures of 200 ms or less (e.g., Wolfe, 1984; J. P. Harris and Calvert, 1989; Raymond and Isaak, 1998). It is hard to envisage that such brief periods of activity could lead to serious depletion of neurotransmitter stores. Second, there is physiologic evidence that even prolonged activation does not cause a decline in output of some visual neurons. Thus, although cortical neurons in the cat certainly do adapt to motion (Hammond et al., 1985) and to flicker (van de Grind, Grüsser, et al., 1972), retinal and geniculate cells do not (van de Grind et al., 1972). This last result implies that there are some visual neurons which do not fatigue with continued activation, suggesting that fatigue is not the reason for the adaptation of others. Third, the time course of recovery from adaptation does not seem to match that expected from neural fatigue. For example, Stromeyer (1978) reports that some visual aftereffects can be elicited days or even weeks after the end of adaptation. It seems that the adapted neurons would have replenished their stores of neurotransmitter within a shorter time than this. Thus the notion of neural fatigue does not seem to offer a total explanation for visual aftereffects, though it might be one component.

Other suggested answers to the puzzle have had two parts. The first has been to point out that the production of aftereffects requires somewhat unusual circumstances, namely, continuous fixation on the same invariant stimulus. This may be common in the laboratory but is rare in the "real world." The second part has been to postulate some mechanism or process which is normally beneficial to its owner but produces perceptual distortions in these rather special situations. The rest of this section considers various suggestions for what this mechanism might be. There are three related themes underlying these ideas. One is that aftereffects are produced by error-correcting mechanisms within the visual system. The second is that aftereffects reflect the visual system's attempts to optimize its coding of the environment. Visual neurons have a restricted dynamic range because their firing rate will not increase above a certain amount, and so there are limits to the range of stimuli which they can code. Adaptation aims to use this limited dynamic range most effectively, by shifting it around to match the range of stimuli in the current environmental conditions. The third theme is that of calibration: how the brain interprets sensory messsages (or the pattern of firing in sensory neurons). Although most of the studies have not involved the MAE, many of the experiments could be redone in, and the theories extended to, the motion domain.

7.2.1 Error-Correcting Accounts

The central idea behind these accounts is that the design of visual system means that it is prone to errors of various kinds, but that these can be removed by suitable correction processes which rely on consistencies and redundancies in visual signals (Andrews, 1964). Some of these errors may arise in the optical system of the eye. For example, it is known that the lens of the eye must produce chromatic aberrrations in the retinal image. Thus the retinal image of a black-white edge will contain color fringes. Similarly, an astigmatism will produce distortions of relative orientation. Other neural changes may arise as the brain ages, and the blood supply to some neurons becomes restricted so that they work less efficiently. Such constant errors can be detected by sampling the visual input over a sufficiently long period. For example, in the image, the color fringes on a black-white edge reverse direction if the edge is rotated through 180 degrees, and the relative orientation of two lines on the retina in astigmat is in changes as the stimulus configuration is rotated. These consistent changes can be removed by suitable neural tuning, provided some assumptions can be made about the environment. For example, the orientation of edges on the retina is probably detected by somehow pooling the activity of a number of neurons tuned to different orientations. Some of these neurons will be strongly excited, others less excited, by the edge, and perceived orientation may be coded by the distribution of their activity. If the blood supply to neurons coding one end of the distribution is reduced, and their firing rate is reduced, the distribution would be skewed, so that perceived orientation would be altered.

How could this potential problem be overcome ? First, an assumption needs to be made about the likely occurrence of different orientations in the world (most simply, that, over a long enough time interval, they are all equally likely). Second, a monitoring device is required which checks whether the activity of individual orientation-sensitive neurons reflects the assumption (so that their time-averaged activity is equal). Third, a mechanism is needed which alters the activity of individual neurons to restore the equality of neuronal activity to the desired state by altering their response characteristics (increasing or decreasing their firing rate to a particular input pattern). In principle, this mechanism would be like a "graphic equalizer" on a sound system, in which different frequency bands in the input signal are processed by different channels whose gain (volume) can be adjusted manually by the listener, to suit his or her own taste and the acoustics of the room. In the visual system, the gains of individual channels (neurons) would be set automatically by a comparison of actual and ideal time-averaged activity. An arrangement of this type has been suggested by Ullman and Schechtman (1982). In normal circumstances, such a mechanism would act to keep its owner's internal representation of

orientation veridical despite unwanted changes (or drift) in individual components of the visual system. However, prolonged viewing of, say, a vertical grating would lead to an excess of activity in vertical neurons, which would be mistaken for a change in their gain, and lead to a reduction in their output. This reduction would manifest itself as a tilt aftereffect (TAE) (e.g., J. P. Harris and Calvert, 1985).

One can give a similar account of the processes underlying contingent aftereffects, such as the McCollough (1965) effect. To obtain this aftereffect, the observer stares for about 10 minutes at a field of vertical black stripes on a red background, alternating every 10 seconds or so with horizontal black stripes on a green background. After this adaptation regimen, black-and-white stripes appear tinged with green when vertical and tinged with pink when horizontal. Anstis (1975) suggests that the system which keeps the coding of, say, color and orientation separate is imperfect, and produces unwanted intermodulation or crosstalk. A possible analogy here might be that of a cable which contains many separate wires, in which activity in one wire (A) can produce spurious activity in a neighboring wire (B). Since this activity in B would always occur when wire A was active, it could be detected and edited out by a suitable filter. During adaptation to colored gratings, then, the brain would treat the correlation in the stimulus between red and vertical as unwanted noise, and turn down the gain of the red mechanism when the vertical mechanism was active. Thus black-and-white vertical edges would produce "anti-red" (or green) activity in the color channels. Such a mechanism would also act to remove the effects of chromatic aberration in the retinal image from the neural image.

Although such accounts clearly explain the basic phenomena of aftereffects, we can ask how well they explain more detailed aspects of the data. They seem to imply that aftereffects should take time to build up and also to decay, since the underlying processes need to sample appropriate aspects of the visual input over time. This fits with Stromeyer's report of the longevity of the McCollough effect, noted above. On the face of it, it does not fit so well with reports of aftereffects from very brief exposures. However, such studies involve a series of short adapting exposures, each followed by presentation of one of a range of test fields in, say, a staircase procedure. Thus, it could be argued that the aftereffects result from the cumulative effect of many short exposures. Error-correcting accounts also imply that recovery from adaptation should not occur simply with the passage of time, but require exposure to a relevant perceptual diet different from that during adaptation. Consistent with this, Spigel (1962a) reported that the MAE can still be obtained if an interval is left between the end of adaptation and presentation of the test field. The MAE still occurs when this interval is longer than the duration of the

MAE when the test field is presented immediately after adaptation. Such "storage" of aftereffects is further evidence against the "neural fatigue" explanation (since presumably neurons would recover from fatigue during the adapt-test interval), but supports an explanation involving sampling of the visual input over time.

7.2.2 Coding Optimization Accounts

7.2.2.1 Redistributing Sensitivity The central problem in explaining selective adaptation is that at first sight it appears to make perception worse. Several studies have tried to show that in fact adaptation can improve some aspects of perception. Barlow, Macleod, et al. (1976) measured various aspects of detection and discrimination of gratings before and after adaptation to gratings which varied in their similarity to the test grating. They found no improvements for the stimulus variables of contrast, spatial frequency, and orientation. However, it may be that the method and, especially, the relationship between the adapting and test stimuli are critical for such improvements to become apparent. Both De Valois (1977) and Tolhurst and Barfield (1978) report increases in sensitivity in detecting gratings after adaptation. However, this did not occur at the adapting spatial frequency (for which sensitivity was reduced, as found in many other studies), but rather when the test grating differed by about two octaves from the adapting grating. De Valois suggests that improved detection arises because neurons tuned to different spatial frequencies inhibit each other. So, in normal circumstances, when one channel is excited by a stimulus, it will not only pass on that information to the rest of the visual system but actively try to prevent other channels (which may be excited to a lesser extent) from doing so. The effect will be to increase the precision of the neural response of the whole system to any stimulus. However, adapting to one spatial frequency reduces not only the output of the most active channel (so that it is less sensitive to its preferred spatial frequency) but also reduces the inhibition which it exerts on other spatial-frequency channels. Thus the latter become more sensitive.

Greenlee and Heitger (1988) measured how different in contrast two successively presented gratings had to be for an observer to discriminate which had the higher contrast. They found that the just noticeable difference (JND) in contrast rose with the absolute contrast of the gratings. The authors then repeated the experiment, preceding every presentation of these test gratings with a period of adaptation to a high-contrast (0.8) grating. Although they still found a dependency of the JND on absolute contrast, the slope of the graph was much shallower than that without adaptation, and the two graphs crossed over at around the value of the

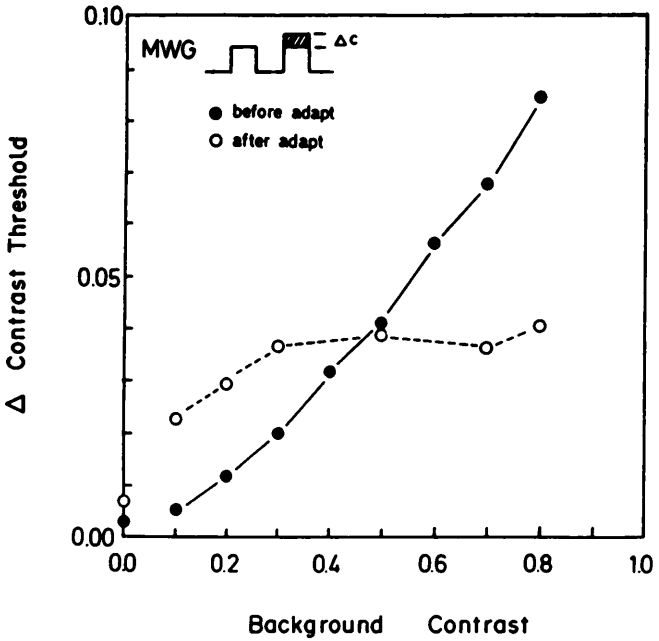

Figure 7.5
Size of the contrast difference (contrast threshold) needed by one subject (MWG) to discriminate which of two gratings had the higher contrast. Values on the abscissa show the background (or standard) contrast on each presentation. Filled symbols show the contrast thresholds before adaptation, open symbols the thresholds after adaptation to a grating of 0.8 contrast. Note that thresholds are lower (discrimination is easier) after adaptation for gratings whose contrast exceeds about 0.5.

adapting contrast (figure 7.5). In other words, for absolute contrasts lower than the adapting contrast, discrimination performance was worse than before adaptation, and for contrasts higher than the adapting contrast, performance was better.

This result was explained as follows. The visual system has a nonlinear response to contrast, so that the plot of perceived or neurally signaled contrast against physical contrast is an S shape rather than a straight line (figure 7.6). At very low or very high contrasts, for which the slope of the graph is very shallow, the change in physical contrast needed to produce a given change in perceived contrast will be large, whereas for medium contrasts (for which the slope of the graph is steeper) this change will be relatively small. Presumably, the JND reflects the size of this change in contrast. Greenlee and Heitger suggest that adaptation to high contrasts shifts the contrast response function, so that some contrasts which previously fell on a steep region now fall on a shallow one, and vice versa.

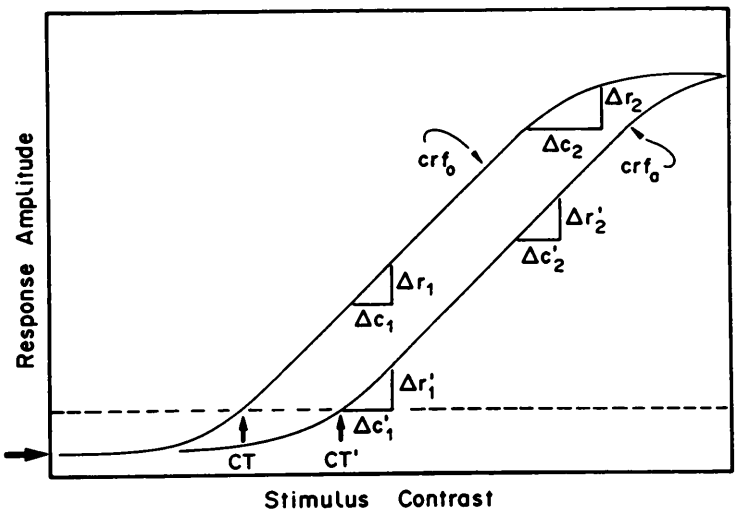

Figure 7.6
Explanation of the result shown in figure 7.5. It is supposed that the output (response amplitude) of the neural mechanisms responding to stimulus contrast follows that of the S-shaped functions shown in the graph. Thus at very low, or very high, contrasts, a larger change of stimulus contrast is needed to produce a given change in response amplitude than at intermediate contrasts. The key idea in the explanation is that adaptation to a high stimulus contrast shifts the response function from position crf_0 to position crf_a. Thus, before adaptation, the change in stimulus contrast (Δc_2) needed to produce a change in response amplitude (Δr_2) is larger than that ($\Delta c'_2$) needed to produce the same change in response amplitude ($\Delta r'_2$) after adaptation. In other words, after adaptation, contrast discrimination thresholds for high-contrast gratings will be smaller than before adaptation, as the data in figure 7.5 suggest.

This would improve discrimination for some contrasts and impair it for others, as their data suggest.

One potential problem with this account is the time course over which the effects operate. One might suppose that the shift in the contrast response function would need to be fast, since, to be useful, it would presumably need to operate within a single glance as one changes fixation from a low- to a high-contrast part of the scene. Greenlee and Heitger provide no evidence that such fast changes can occur, though there is certainly evidence, noted above, that aftereffects can result from a series of very brief stimulus presentations.

There is a distinction to be made between the models of De Valois and of Tolhurst and Barfield, on the one hand, and of Greenlee and Heitger, on the other. The results of all three studies show a redistribution of sensitivity to a particular stimulus attribute produced by adaptation to some value of that attribute. Thus, after adaptation, observers are more sensitive to some value of the attribute and less sensitive to others. For

Greenlee and Heitger, this occurs because adaptation acts to readjust the nonoptimized coding system within a single channel to the presently prevailing visual diet. Although the other authors do not explicitly discuss its functional significance, in their accounts adaptation appears to act by disrupting a system already optimized by mutual inhibition between separate neural channels.

7.2.2.2 Decorrelation These ideas about aftereffects reflecting mechanisms that optimize neural coding have been extended by Barlow (1990). He suggests that the cortex is a device for detecting the occurrence of novel events, and changes its own organization on the basis of correlations between different features of the environment. Barlow first considers contingent aftereffects, using the example of the aftereffect of color contingent on orientation (the McCollough effect—see above). He explains this effect as follows, with reference to figure 7.7a–d. In figure 7.7a, ψ_A on the vertical axis and ψ_B on the horizontal axis represent two perceptual variables, each capable of discriminating only four values (say blue, green, yellow, red; and horizontal, left oblique, vertical, right oblique), which depend on the values of two physical variables, A and B (color and orientation). The points on the graph show how which combinations of values of the two physical variables have occurred over some period of time. The variables are uncorrelated, so that all colors are about equally likely to have occurred with a particular orientation (and vice versa). Moreover, the combination of perceptual variables represents the combination of physical variables well, since all sixteen regions of the graph have some points in them. In figure 7.7b, the physical variables are correlated, and so particular colors occur only in combination with particular orientations (as they do during McCollough adaptation). In this case, many cells in the graph have no points in them because those combinations of color and orientation never occur, so that the coding of environmental events is inefficient, with only seven out of sixteen regions containing points. The solution is, in effect, to rotate the axes of the graph, so that the perceptual dimensions represent the physical dimensions more efficiently, since all cells now have some points within them, as shown in figure 7.7c. When this "oblique" graph paper is stretched (figure 7.7d), so that the perceptual axes are orthogonal again, the axes for the physical dimensions are now oblique. Thus the physical variable A, which was originally plotted vertically, now has a negative component on the perceptual axis plotted horizontally, giving, say, the negative contingent aftereffect of color found by McCollough.

Barlow suggests that the lesson to be drawn from such aftereffects is that "perceptions are intended to occur independently, and define independent axes in perceptual space." When stimulus dimensions are

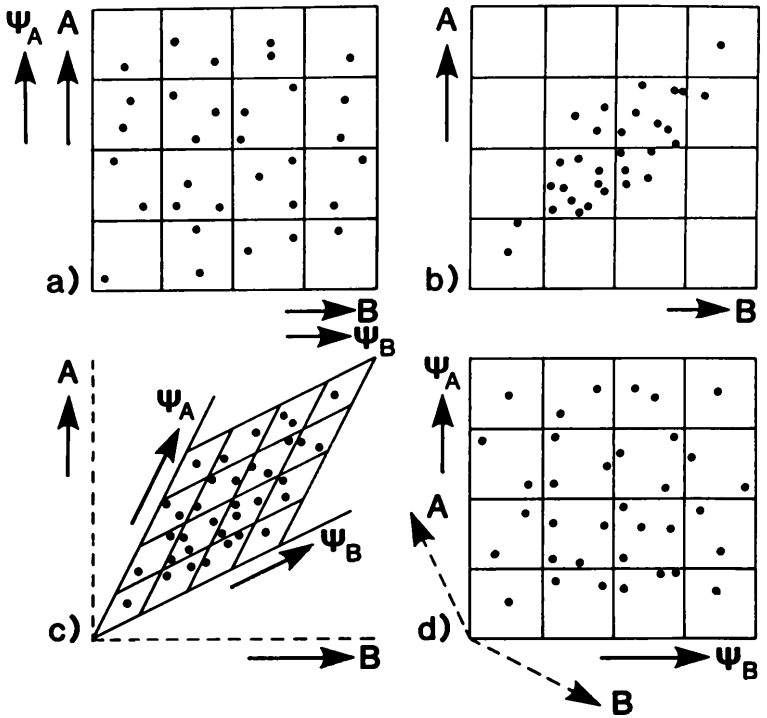

Figure 7.7
Two perceptual variables, Ψ_A and Ψ_B, each capable of discriminating only four levels, depend principally on two physical variables, A and B. In figure 7.7a, the two physical variables are uncorrelated, and the scatter diagram of joint occurrences shows that they fill all the cells in the diagram. In figure 7.7b, on the other hand, the physical variables are correlated. If the perceptual variables were simply proportional to A and B, the coding scheme would be inefficient, because some joint values do not occur and many of the cells are empty. The solution is to rotate the perceptual axes (figure 7.7c), so that the grid fits the joint values of A and B which actually occur, and all sixteen cells of the grid are filled. This can be seen when the grid is replotted with Ψ_A and Ψ_B as orthogonal axes instead of A and B (figure 7.7d). Now the "perceptual surface" is fully covered by the joint values. The rotation of the perceptual axes is thought to be the equivalent of adaptation to McCollough stimuli. In figure 7.7d, the axis of the physical variable A has a backward tilt, implying a negative component on the percept Ψ_A, as would be expected from the occurrence of the McCollough effect.

artificially linked, as in McCollough adaptation, then a repulsive force develops between the perceptions which spreads out the cluster of responses in perceptual space. This "repulsive force" might be mutual inhibition between the neural systems underlying each separate percept. The result of a large number of such repulsive forces, built up by the correlations and redundancies of a particular visual environment, would be a visual system which tended to produce little activation in response to familar combinations of stimulus attributes, but a lot of activation to novel stimuli.

Although this idea, that aftereffects reflect mechanisms which "decorrelate" different stimulus attributes, applies most obviously to contingent aftereffects, Barlow extends it to simple aftereffects, such as the MAE. He argues that these occur because adapting stimuli are large enough to cover the receptive fields of many neurons. During continuous motion, many neurons will be active simultaneously, and so will inhibit each others' activity. When a stationary test field is viewed, the inhibition will be sustained and so the appearance of reversed motion will be produced. On this view, aftereffects result because the adapting stimuli produce correlated activity in a group of neurons, rather than activity in single neurons.

Some physiologic evidence in support of this idea has been provided by Carandini, Barlow, et al. (1997). They measured the responses of cells in the primary visual cortex of the monkey to each of two gratings, oriented at 90 degrees to each other, and to the compound stimulus (or plaid) formed by presenting the gratings simultaneously. One of the gratings (G1) was oriented in the preferred orientation of the cortical cell, while the other (G2) was oriented at right angles to G1. The authors varied the contrast of the gratings and of the plaid, and measured the responses to a range of stimulus contrasts, after continuous stimulation of the cell by (or adaptation to) high-contrast versions of these stimuli. An important aspect of the findings was as follows. The responses of the cells to G2 (always orthogonal to the cell's preferred orientation) were negligible. Thus, if it were simply the physical characteristics of the stimulus which governed the adaptation of the cell, adapting to G1 alone should have the same effect on tests of G1 alone and G1 + G2 as adapting to G1 + G2. However, this was not so. Adaptation affects were much larger when the adapting and test stimuli, whether plaid or grating, were the same than when they were different. In other words, although G2 itself was an ineffective adaptor, the compound G1 + G2 had adapting effects for a G1 + G2 test which were greater than the effects for adapting to G1 alone. This result is consistent with the authors' suggestion that the cells were adapting to the contingent occurrence of G1 and G2, and is physiologic support for Barlow's ideas about the role of selective adaptation.

7.2.3 Recalibration

A central problem for the brain must be the interpretation of sensory activity. Unlike the laboratory scientist, who can measure with a ruler, for example, how many centimeters a lever must be moved to produce a given change of voltage in some apparatus, the brain has no metric of the external world which is independent of its own activity. Calibration of sensory messages can only be done on the basis of assumptions about the nature of the world. Examples of such assumptions might be that, averaged over a long enough time period, all orientations or directions of motion are equally likely to occur in any region of the retinal image. These ideas were touched on earlier in discussing error-correcting devices. Like Andrews, J. J. Gibson (1937) noted that the brain has a potential problem in keeping the physical and phenomenological worlds in correspondence. He pointed out that many sensory dimensions have a norm or null point. For example, stationarity (or absence of motion) can be thought of as a null or midpoint on a continuum running from, say, fast motion to the left through to fast motion to the right (or as the midpoint of a two-dimensional space). Gibson's account essentially suggests that this null point, or norm, is somehow calculated by the brain from the stream of sensory information about that particular stimulus dimension. As he put it, there is "a tendency for sensory activity to become normal, standard or neutral" (Gibson, 1937, p. 226). Put another way, his view is that the value of the null point of a sensory dimension is not wired into the brain, but represents, say, the average activity on that dimension over the recent past. Adaptation biases that activity, and so shifts the null point. This means that after adaptation to, say, movement to the left, stimuli which fall on the old null point (stationary) now no longer do so, but appear to move to the right. This idea suggests that the brain must continually recalibrate its inputs to optimize the correspondence between the external world and its internal visual representation.

If this view of perception is correct, then interpreting sensory messages must involve a comparison of the present sensory state with some longer-term measure of sensory activity, since the latter provides the only reliable reference. In their discussions of the functions of the processes underlying the McCollough effect and related contingent aftereffects, both Dodwell and Humphrey (1990) and Durgin and Proffitt (1996) point out that this idea is essentially that behind Helson's (1948, 1964) adaptation level theory. In Dodwell and Humphrey's words, "The most important idea in adaptation level theory is that the 'neutral point' (adaptation level), in some sense the 'center' of psychophysical judgements, is a weighted average of the set of stimuli so far presented" (p. 79).

For moving stimuli in the real world, then, stationarity (lack of retinal motion, or the neutral point of the "motion scale") would be the time-

averaged activity of motion-sensitive neurons. The brain would not have to make strong assumptions about the consistency and reliability of its own internal machinery, but rather assume that the world was consistent, and that this consistency provides a potentially reliable reference. There have been various suggestions about what the brain does with this reference once it has been extracted. Dodwell and Humphrey suggest that an error-correcting device (like that of Andrews) operates to change the values ascribed to particular patterns of sensory activity, in order to maintain a correspondence between the world and its internal representation. Durgin and Proffitt prefer the ideas underlying Barlow's model: the reference can be used in a system giving efficient sensory coding, while at the same time highlighting novel sensory events.

The problem with a system which relies on long-term statistical properties of the input is its vulnerability to atypical short- and medium-term changes. Inevitably, these will bias the reference and so change the way in which subsequent sensory events are interpreted, as initially suggested by Gibson.

All the above accounts have in common that they discuss relatively local processes, confined to adapted areas of the retina. However, some types of adaptation can produce more global changes in perception in which the subject's entire frame of reference may be altered. Much of Gibson's experimental work on adaptation concerned the TAE (Gibson and Radner, 1937). One important aspect of this work was the demonstration that the vertical and horizontal axes of visual space could be linked in some way. So, after adapting to a line slightly off vertical, a small aftereffect was found on a horizontal test line. Gibson called this the "indirect" effect, to distinguish it from the "direct" effect on a vertical test line. This indirect effect implies that adaptation to a line close to vertical can distort the whole visual frame of reference, rather than simply affect the perception of stimuli which are similar to the adapting stimuli. Morant and Harris (1965) showed that in addition to this "global" effect on the visual frame of reference, there is also a "local" effect, which is confined to test stimuli similar to the adapting stimuli. Presumably, the local and global effects of adaptation to tilt (and by implication to motion also) reflect processing at different levels of visual analysis.

As noted earlier, adaptation is known to occur at several cortical sites, and the local and more global effects of adaptation may be the perceptual correlates of activity in these different anatomical sites. Wenderoth (e.g., Wenderoth and Johnstone, 1987) has suggested that different effects originate in different cortical areas, local effects perhaps in V1, more global effects in extrastriate cortical areas, such as V4 or MT. For example, MT may be involved in the perception of the speed and direction of a drifting plaid (and the MAE which results from it), whereas mechanisms in

V1 may respond to the component gratings of which the plaid is formed or the "blobs" of luminance where the gratings cross (see, e.g., Wenderoth et al., 1994; and chapter 5). A similar account can be given for the direct and indirect components of the tilt illusion (Wenderoth and Johnstone, 1987). More recently, Wiesenfelder and Blake (1992) have reported evidence for multiple sites of adaptation in the MAE, based on the use of binocular rivalry, in which a stimulus presented to one eye can suppress the information from the other eye. It had already been shown that the strength of monocular MAEs from adaptation of the same eye is unaffected by the presence of a rivalrous stimulus seen by the other eye, which suppressed the perception of adapting motion (Lehmkuhle and Fox, 1975). This suggests that the MAE is generated before the site at which visual information is blocked by binocular rivalry suppression. When, however, Wiesenfelder and Blake looked at the effect of a binocularly suppressed test field on storage of the MAE, they found a different picture. If, after monocular adaptation, the presentation of the test field to the adapted eye is delayed until the MAE obtained with immediate presentation of the test field would have decayed away, an MAE can still be obtained. It turns out that rivalrous suppression of an immediately presented test field permits this storage of the MAE, just as physically removing the test field would. This suggests that storage and decay of the MAE must be mediated at least in part by processes which lie after the site of rivalry suppression. It is tempting to attribute the presuppression adaptation stage to changes in motion sensors, and the postsuppression storage stage to activity in the integrator or higher levels.

MAEs do not store perfectly, in the sense that stored MAEs are weaker than MAEs measured immediately after the same adaptation regimen, as noted by Wiesenfelder and Blake, as well as by other workers. Thus some decay of the MAE occurs even in the absence of any test field.Wolfe and O'Connell (1986) measured the TAEs produced by varying periods of adaptation. They found that the TAE from 2 minutes of adaptation decayed away within 4 minutes, whereas the TAE from 4 minutes of adaptation could still be measured after 2 weeks, even though, at the end of adaptation, the TAEs from the two periods of adaptation were of similar magnitude. The authors suggested that the fast-decaying component of adaptation occurs in "broadly-tuned channels," and perhaps reflects neurotransmitter depletion (arguably, neural fatigue). On the other hand, the longer-lasting component was thought to reflect a change in the activity of "labelled-lines," which detect ratios of activity between the broadly tuned channels. Although Wolfe and O'Connell invoke neural fatigue in their explanation of the TAE, the data do not force such an explanation upon us. The best evidence for neural fatigue seems to be the apparent decay of the MAE in storage experiments in the absence of a

test field. However, the walls of an experimental laboratory or the surfaces of experimental apparatus have a microtexture, as well as the dark field produced by closing the eyes. Thus the visual system is being presented during the storage interval with information about stationary patterns, as would be required by accounts such as recalibration.

Although some of the detail seems open to dispute, this suggestion of multiple sites of adaptation seems to fit well with the data from the binocular rivalry experiments. Thus imperfections of storage of MAEs would result from changes in presuppression opponent-energy sensors, whereas the stored component of MAEs would reflect changes of integration.

So far, it has been suggested that visual calibration takes place relative to the statistical properties, over time, of the retinal image alone. However, there are other sources of information which could, in principle, affect the interpretation of visual activity, namely, vestibular and proprioceptive (and perhaps auditory) information, and the corollary discharges associated with motor activity. L. R. Harris et al. (1981) suggested that the MAE might result from a process which calibrates the relationships between different sensory inputs. They pointed out that the most common cause of retinal motion is not motion of the environment but motion of the observer. Thus, for example, the expanding optical flow on the retina produced by forward locomotion is normally accompanied by correlated signals from the vestibular system. To check the idea that the MAE might result from an unusual mismatch between vestibular and retinal signals, they placed the adapting display, and in some conditions the observer, on a movable trolley. Sinusoidal-to-and-fro motion of the trolley was converted via the voltage across a potentiometer into expansion and contraction of a field of dots on an oscilloscope screen. The authors found a strong contracting MAE , from retinal expansion without observer motion, but this was markedly reduced when the observer moved with the display, as the intersensory recalibration hypothesis suggests. However, one might have expected a similar reduction in strength of the MAE resulting from retinal expansion due to backward motion, and this was not found: the expansion MAE was only slightly reduced when the observer moved backward with the display. Nor was the MAE enhanced when the direction of motion on the retina and the direction of observer motion were put into conflict.

Despite these apparent discrepancies within their experiment, however, there is other evidence for the kind of intersensory recalibration suggested by these authors. An experiment complementary to that of L. R. Harris et al. (again changing the usual relationship between retinal and vestibular signals) would be to have the observer move during adaptation, but to keep the retinal image motionless. After jogging on a tread-

mill for 10 minutes, subjects report a sensation, when walking normally on solid ground, of moving at an accelerated rate (Pelah and Barlow, 1996). The authors, who describe other related illusions, conclude that disturbing the normal relationship between self-induced motion and expected sensory input leads to a recalibration of the relationship between optic flow, vestibular signals, and movements of the legs. One way to describe this illusion is as an MAE produced by the absence of visual motion where such motion would normally occur.

It is not yet certain whether the site of these intersensory MAEs is the integration level described in the previous section. If so, they should show the same patterns of binocularity and spatial tuning as other MAEs which are thought to reside there. Such experiments have yet to be done, though Pelah and Barlow note that, after adaptation with a textured wall on one side, their effect was stronger when walking with a wall on that side rather than the other. Without such evidence, it is not clear whether one needs to postulate a third, higher, level of motion adaptation, at which visual and nonvisual information is integrated. Whatever the answer, it seems that the mechanisms underlying these global (integrator) effects may save the same functions as those underlying the local (sensor) effects. That is, drift or optical errors mean that, say, the perceived vertical or the perceived stationarity of the whole visual field needs to be continuously recalibrated; or the range of possible orientations or directions of motion need to be redistributed across the available mechanisms to suit particular visual environments.

7.2.4 Which Account Is Best?

One difficulty in deciding between error-correcting, coding optimization, and recalibration accounts of motion adaptation is that they appear to make very similar predictions. They allinvolve monitoring activity in visual mechanisms over time, suggesting that aftereffects should build up relatively slowly, and also decay slowly, since the visual system needs time to take account of the change of visual (or other perceptual) diet between adaptation and testing. They all appear to predict storage of aftereffects between adaptation and presentation of the test field, since it is an alteration of visual or other input, not simply the passage of time, which is needed to readjust the underlying mechanisms. However, there are situations in which the three accounts seem to make different predictions. For example, the error-correcting account implies that MAEs should be stronger the more characteristics are shared by the adapting and test fields, since it is a subset of motion-sensitive mechanisms which would be affected by the test field. On the recalibration account, however, evidence of, for example, absence of movement could come from a test field with spatial charateristics very different from those of the adapting field.

Although the error-correcting, optimization, and recalibration accounts have been presented as alternatives, they do not exclude one another. The same kind of mechanism or process within the visual system could fulfill all these roles. To illustrate this point, consider the human nose. Although one can ask whether the function of the nose concerns respiration or olfaction, the answer is clearly "both." Indeed, it is just because breathing through the nose produces a regular flow of air over the nasal membranes that it is a prime site for olfactory receptors. Thus it may be that "self-tuning" devices of the kind outlined here can fulfill all these "housekeeping" functions in vision. One possibility is that adaptation in opponent-energy sensors is best thought of as error correction, whereas that in higher-order integrators reflects optimization and calibration processes.

7.2.5 Conclusions

Early views that selective adaptation reflects neural satiation or fatigue are probably inadequate, since they are not consistent with evidence on the buildup and decay of aftereffects, or the evidence that some visual neurons do not fatigue with continuous stimulation. Alternative accounts (error correction, coding optimization, and recalibration) fit the evidence better, and present evidence does not decisively favor one of these over the others. They are not mutually exclusive, and all may be correct.

Adaptation occurs at several cortical sites, and this may be reflected in a range of motion, tilt, and other aftereffects. For example, there seem to be two types of TAE, one to do with local orientation processing, the other with the more global frame of reference ("perceived vertical"). There seem to be analogous MAEs.

MAEs can result from the interaction of visual and nonvisual signals. It is not yet clear at which level of motion analysis this interaction occurs.

7.3 General Conclusions

A strong theme to emerge from section 7.1 was the need for models of motion analysis containing several layers of processing, with adaptation arising at each layer. Without computational modeling, it is not clear just how well such models can account for the detailed properties of many MAE phenomena reported in this book. However, the recent emergence of new stimulus paradigms in MAE research has provided new data against which to test computational models, so the way is open for significant theoretical advances in the near future. New ideas on the significance of adaptation, described in section 7.2, hint at the functional logic behind multiple adaptation sites in motion processing. Short-term imbalances between excitation and inhibition are highly significant, because they indi-

cate directional bias in the image, either locally if they arise from sensor responses, or globally if they arise from recurrent connections between integrators. Selective adaptation may serve to ensure that, over a longer time scale, excitation and inhibition in different layers tend to balance out.

Note

1. Second-order motion sensors can be constructed using the same sequence of processing as used by first-order motion sensors, with one additional operation: a nonlinear transformation (e.g., rectification) is applied to the signal before it is subjected to motion-energy analysis, to convert texture modulation into "intensity" modulation in the neural image. There is good evidence for the existence of both kinds of detector, and Wilson and Kim (1994) accordingly assumed that both first-order and second-order opponent-energy responses sum their responses in the pattern layer of the model.