9 Neuronal Correlates of Perceptual Organization in the Primate Visual System

Thomas D. Albright
Howard Hughes Medical Institute
Salk Institute for Biological Studies
Lisa J. Croner
Robert O. Duncan
Gene R. Stoner
Salk Institute for Biological Studies

The primate visual system has been studied by neurobiologists for nearly 4 decades. As a result, we know a great deal about how neurons at early levels of visual processing represent basic features of the retinal image, such as the orientation of a contour or the color of a patch of light. Nevertheless, we know very little about how such cells give rise to our perceptual experience of the world. One reason for this limited knowledge is that perception is context dependent. Just as the meaning of a word depends on the sentence in which the word is embedded, the perceptual interpretation of a retinal image feature depends on the spatial and temporal context in which the feature appears. Until recently, contextual manipulations have been largely excluded from studies of the response properties of visual neurons, primarily as a matter of investigative convenience.

To appreciate the importance of the contextual influence on perception, consider Fig. 9.1. Panels A–C each contain an identical, horizontally oriented bar. The perceptual interpretations of the three bars are very different: Human observers typically report that the bar in Panel A appears to be a region of overlap between two light gray rectangular surfaces, one of which is transparent. The bar in Panel B appears as a variation in surface reflectance, as though a dark gray stripe has been painted horizontally across a larger light gray planar surface. The bar in Panel C appears to be a shaded portion of a larger object that is folded in three dimensions.

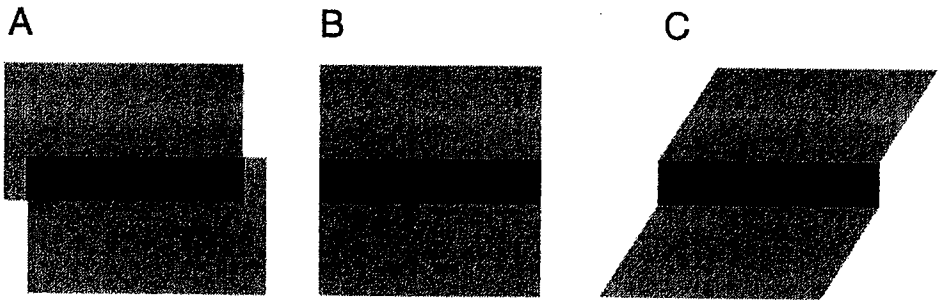

FIG. 9.1. Illustration of the influence of visual context on perception. Each of the three images displayed here contains a horizontal dark gray rectangle. Although these rectangles are physically identical, the surrounding features (the contexts) differ in the three images. As a result, the rectangle is perceptually attributed to different environmental causes in the three instances. (A) The rectangle appears to result from the overlap of two surfaces, one of which is transparent (e.g., a piece of tinted glass). (B) The rectangle appears to result from a variation in surface reflectance (e.g., a stripe painted across a large flat canvas). (C) The rectangle appears to result from partial shading of a surface (i.e., variation in the angle of the surface with respect to the source of illumination). These markedly different perceptual interpretations argue for the existence of different neuronal representations of the rectangle. These representations can only be identified in neurophysiological experiments if appropriate contextual cues are used for visual stimulation.

The different perceptual interpretations of the dark gray bars in Fig. 9.1, Panels A–C, follow from the different spatial contexts in which the bars appear. Such contextual influences, which are ubiquitous in normal visual experience, raise two important questions for neurobiologists: (1) At what stage of visual processing is contextual information incorporated to achieve neuronal representations of things perceived (scene-based representations), rather than local features of the retinal image (image-based representations)? (2) What neuronal mechanisms underlie the transformation from image-based to scene-based representations?

Our laboratory spent much of the past decade addressing these questions, primarily in the domain of visual motion processing (for review see Albright, 1993; Croner & Albright, 1999b; Stoner & Albright, 1993). In this review, we summarize two recent experiments from our laboratory. These experiments address the phenomenology and neuronal bases of two key elements of perceptual organization: feature interpretation and feature grouping.

FEATURE INTERPRETATION

When viewing any natural scene, the visual system interprets each feature—such as a corner, edge, patch, and so on—as belonging to a particular object. Fig. 9.2 illustrates the relevance of feature interpretation to motion processing. The scene in

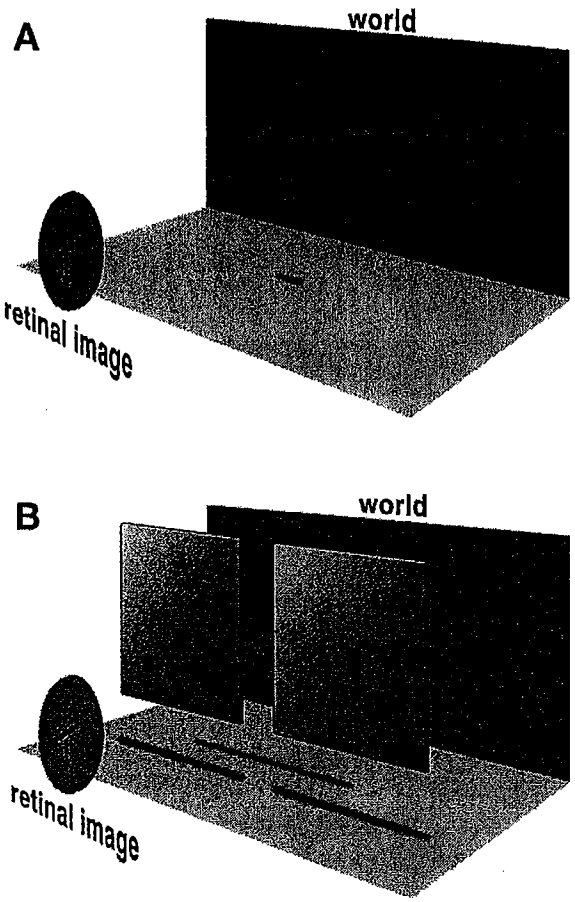

FIG. 9.2. Illustration of the contextual dependence of motion perception. Two simple visual scenes are shown, along with the retinal image motions that they give rise to. (A) This visual scene contains a single moving object (an elongated bar), which is oriented obliquely and moving directly down. The retinal image motion directly reflects the object motion. (B) This visual scene contains a single obliquely oriented bar, but in this case the bar is moving to the lower right (in a direction orthogonal to its orientation). The scene also contains two static opaque occluders, which block the observer's view of the two ends of the moving bar. The resulting retinal image motion in Panel B is identical to that caused by the scene in Panel A, although the visual scene motions are clearly different. The different contexts present in the two scenes—no surrounding features in Panel A, occluding panels present in Panel B—allow the different visual scenes to be perceived veridically.

Fig. 9.2A contains a single moving object—a sticklike figure—oriented obliquely and moving directly downward. The left-hand side of the figure illustrates the retinal image rendered by this simple scene. The scene in Fig. 9.2B also contains a single moving object—once again, a sticklike figure—obliquely oriented but moving toward the lower right corner of the scene. Unlike the scene in Fig. 9.2A, this scene also contains two opaque surfaces, each occluding one end of the moving

object. The resulting retinal image of the moving object, shown at left, is an obliquely oriented line moving directly downward. Importantly, the retinal image motion rendered by the bar in the second scene (Fig. 9.2B) is physically identical to that rendered by the bar in the first scene (Fig. 9.2A), although the object motions are clearly different.

Human observers perceive the different object motions correctly in the two cases, and they do so because of the two different spatial contexts (with occluding panels in Fig. 9.2B; without panels in Fig. 9.2A), in which the image motions appear. In particular, the two contexts lead to different interpretations of the terminations of the moving line in the retinal image (Shimojo, Silverman, & Nakayama, 1989). In Fig. 9.2A, the line terminators of the retinal image are interpreted as the ends of an elongated object in the visual scene (i.e., they are intrinsic features of that object). As such, their downward motions reflect the motion of the object itself. By contrast, the line terminators of the retinal image in Fig. 9.2B are accidents of occlusion and thus extrinsic to the moving object. Unlike the intrinsic features of Fig. 9.2A, the downward retinal image motions of the extrinsic features in Fig. 9.2B do not reflect the motion of the object and hence do not figure into the computation of object motion. Thus, the manner in which the terminator features are interpreted determines how the retinal image motions are perceived.

These perceptual phenomena prompt us to ask whether visual neurons respond differently to the motions represented in Fig. 9.2A versus Fig. 9.2B, thereby encoding object motion rather than image motion. Experiments conducted over the past 40 years yielded a wealth of information regarding the response properties of motion-sensitive neurons in the primate visual system. Consider, for example, the responses illustrated in Fig. 9.3. These data, which were published by Hubel and Wiesel in 1968, document a property known as directional selectivity, which is now known to be common among visual cortical neurons (Albright, 1993). The data illustrated were recorded from a single neuron in the primary visual cortex (area V1) of a rhesus monkey. The visual stimulus was an oriented bar of light (oriented lines in left column of Fig. 9.3) that was swept through the receptive field (dashed rectangles) of the neuron under study. Neuronal responses to these stimuli are shown on the right-hand side of the figure as spike trains, in which action potentials (vertical lines) are plotted as a function of time. The stimulus that elicited the largest number of action potentials was oriented diagonally and drifted to the upper right through the receptive field. Response magnitude waned as the motion deviated from this preferred direction.

Observations of this sort have been made repeatedly since the discovery of neuronal directional selectivity, and they have revealed much about the way in which cortical neurons encode image motion. Do these findings enable us to predict how motion sensitive neurons will behave when presented with stimuli such as those in Fig. 9.2 or under other stimulus conditions that approximate the richness of normal perceptual experience? If such cells were sensitive only to light within their

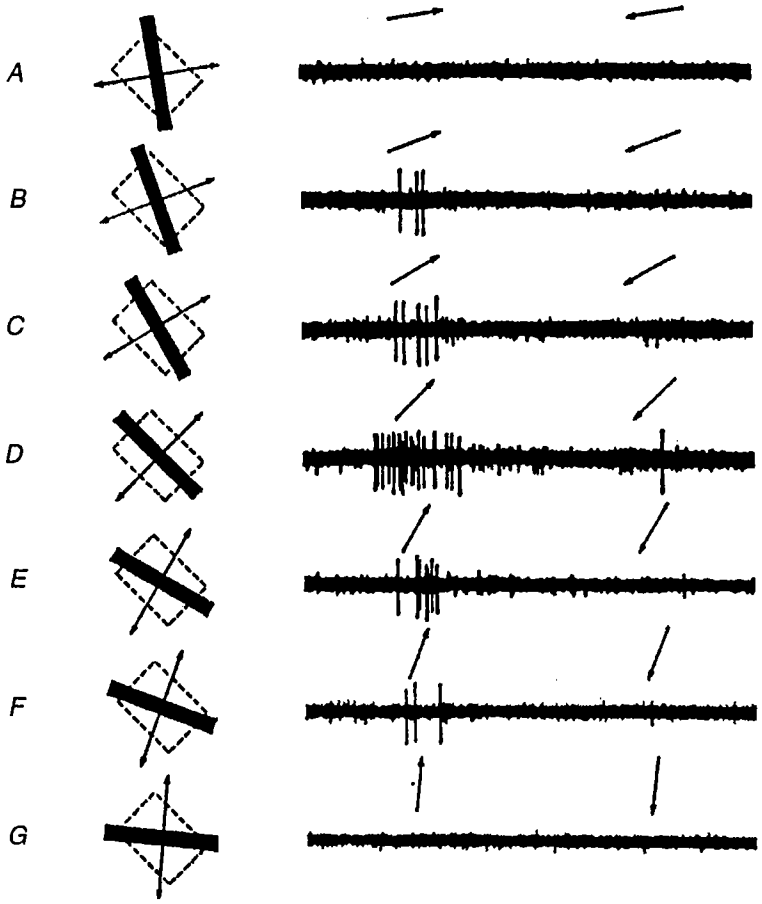

FIG. 9.3. Directional selectivity of one neuron, as first observed by Hubel and Wiesel (1968) in primary visual cortex (area V1) of rhesus monkey. The neuronal receptive field is indicated by broken rectangles in the left column. The visual stimulus was a bar of light moved in each of 14 directions (rows A–G, opposing directions indicated by arrows) through the receptive field. Recorded traces of cellular activity are shown at right, where the horizontal axis represents time (2s/trace) and each vertical line represents an action potential. This neuron responds most strongly to motion up and to the right (row D). From "Receptive Fields and Functional Architecture of Monkey Striate Cortex," by D. H. Hubel and T. N. Wiesel, 1968, *Journal of Physiology, 195,* p. 214. Copyright 1968 by The Physiological Society. Reprinted with permission.

receptive fields, their responses would be a simple function of the receptive field profile and the spatiotemporal pattern of light falling in the receptive field. Evidence suggests, on the contrary, that some modulatory effects arise from stimulation outside of the classically defined receptive field (CRF). It is therefore impossible to predict how these neurons will respond to a stimulus that is more complex than that used to characterize its CRF.

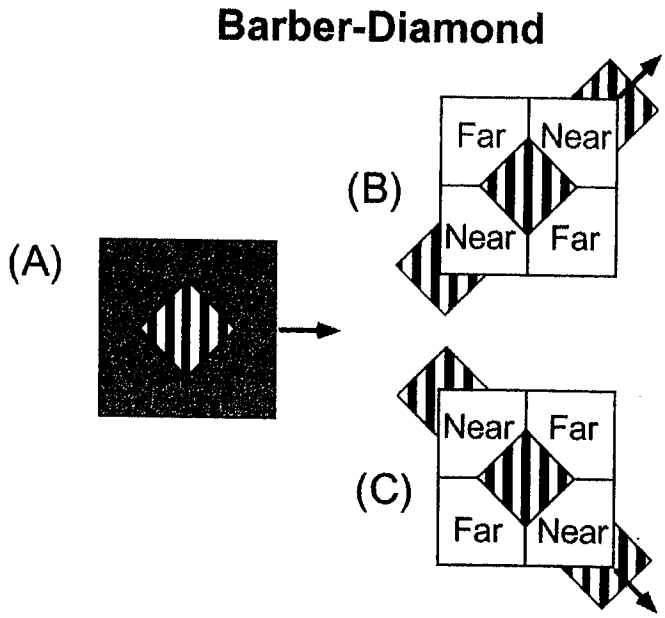

FIG. 9.4. Schematic depiction of barber-diamond stimuli used to study the influence of context on perceived motion and on its neuronal correlates (a demonstration can be seen at http://www.cnl.salk.edu/~gene/). (A) Stimuli consisted of a moving square-wave grating framed by a static, diamond-shaped aperture. The grating itself was placed in the plane of occular fixation, and four textured panels that defined the aperture were independently positioned in depth using binocular disparity cues. The grating was moved either leftward or rightward on each trial. Each of the four stimulus conditions used was a unique conjunction of direction of grating motion (right vs. left) and depth-ordering configuration. The two conditions illustrated (B and C) were created using rightward moving gratings; two additional conditions (not shown) were created using leftward moving gratings. Near and Far identify the depth ordering of the textured panels relative to the plane of the grating. For the condition shown in B, the upper right and lower left panels were placed in the near depth plane, whereas the upper left and lower right panels were in the far depth plane. The line terminators formed at the boundary of the near surfaces and the grating are classified as extrinsic features resulting from occlusion, and the grating appears to extend behind the near surface. (Note that gray stripes are not part of the stimulus and are used solely to illustrate perceptual completion of the partially occluded grating.) Conversely, line terminators formed at the boundary of the far surfaces and the grating are classified as intrinsic (i.e., they appear to result from the physical termination of the surface on which the stripes are painted). As a result of this depth-ordering manipulation and ensuing feature interpretation, observers typically perceive the moving grating in B as belonging to a surface that slides behind the near panels and across the far panels (i.e., to the upper right). This direction is identified with motions of intrinsic terminators. The condition shown in C contains the same rightward grating motion but employs the depth-ordering configuration that is complementary to B. In this case, observers

Most neurophysiological studies to date have not manipulated visual context; hence, there are no existing data that enable us to determine whether visual neurons will respond similarly or differently to the two configurations in Fig. 9.2. We hypothesize, however, that the visual system first extracts simple features within an image and then integrates the neural representations of these features—using context as a guide—to create a new scene-based representation that reflects the structure and properties of the scene. If this were the case, then there should be a population of neurons whose responses are sensitive to the context surrounding their CRFs. We undertook studies to locate such neurons, hypothesizing that they would (1) be differentially sensitive to the motions of intrinsic vs. extrinsic image features, and (2) represent the motions of visual scene elements (i.e., objects and their surfaces) rather than simply the motions of retinal image features.

Barber-Diamond Stimuli

To identify neurons sensitive to visual scene motion, we developed visual stimuli in which a simple contextual manipulation dramatically affects feature interpretation and motion perception (Duncan, Albright, & Stoner, 2000). That stimulus, which we have termed the barber-diamond because of its close relationship to the classic barber-pole illusion (Wallach, 1935), is illustrated schematically in Fig. 9.4. As shown in Fig. 9.4A, the barber-diamond is composed of a pattern of vertically oriented stripes viewed through a diamond-shaped window. The stripes can be moved leftward or rightward within the stationary window. The window is surrounded by a randomly textured field composed of four adjacent panels that could be independently placed in either near or far depth planes (relative to the plane of ocular fixation) using binocular disparity cues.

In our experiments, we manipulated two independent variables: the direction of motion of the stripes (rightward or leftward) and the relative placement of the textured panels in different depth planes. Two depth configurations, shown in Fig. 9.4B and 9.4C, were used. In one case (Fig. 9.4B), the upper left and lower right textured panels were placed in a far depth plane relative to the stripes/fixation point, and the upper right and lower left panels were placed in a near depth plane. In the other case (Fig. 9.4C), the upper left and lower right textured panels were placed in a near depth plane, and the upper right and lower left panels were placed in a far depth plane. The conjunction of two directions of stripe motion and the two depth configurations yielded four distinct stimulus conditions.

FIG. 9.4. (*continued*) typically perceive the grating belonging to a surface that is drifting to the lower right. See text for details. From "Occlusion and the Interpretation of Visual Motion: Perceptual and Neuronal Effects of Context," by R. O. Duncan, T. D. Albright, and G. R. Stoner, 2000, *Journal of Neuroscience, 20,* p. 5890. Copyright 2000 by The Society for Neuroscience.

The two different depth configurations were expected to elicit different intrinsic versus extrinsic feature interpretations for the terminators of the vertical stripes. Specifically, line terminators lying adjacent to far panels were expected to be interpreted as intrinsic to the stripe because the depth edge provides evidence that these terminators lie at the edge of a striped surface. By contrast, line terminators lying adjacent to near panels were expected to be interpreted as extrinsic features because the depth edge here suggests that their termination by occlusion is a more likely possibility. Following this logic outlined, we predicted that motion of the vertical stripes—which was, we emphasize, always physically leftward or rightward along the horizontal axis—would be perceived to follow the motions of the intrinsic terminators (thereby ignoring extrinsic terminator motion). Thus, for the example illustrated in Fig. 9.4B, perceived motion should be toward the upper right corner. Likewise, for the example shown in Fig. 9.4C, perceived motion should be toward the lower right corner.

Feature Interpretation Influences Perceived Motion

We first investigated how the barber-diamond stimuli (Duncan et al., 2000) were perceived by human subjects. For each trial, the subject fixated a small 0-disparity fixation target at the center of the stimulus, while the vertical bars of the barber-diamond moved for 1.5s. After each trial, the subject indicated the perceived direction of motion by orienting a bar displayed on the display monitor. The results for each of the four stimulus conditions are plotted in Fig. 9.5 as frequency distributions of directional reports. Data are summed over all seven subjects studied. Subjects overwhelmingly reported that perceived direction of motion was in the direction of intrinsic terminator motion. These findings support our hypothesis that feature interpretation, derived from contextual cues unrelated to motion, has a marked influence on perceived motion. A demonstration of this perceptual effect can be seen at http://www.cnl.salk.edu/~gene/.

Cortical Motion-Sensitive Neurons Encode Object Motion Using Intrinsic Features

Having confirmed that perceived motion direction depends on contextual cues for depth ordering, we next sought to identify the neuronal bases of these effects. Fig. 9.6A provides a partial summary of the organization of the visual cortex in nonhuman primates (rhesus monkeys), based on a large number of neurophysiological, neuroanatomical, and neuropsychological studies conducted over the past 30 years. Some of the major sulci in the posterior portion of the brain have been partially opened to illustrate the visual areas lying within these sulci. Approximate boundaries between visual areas are identified by broken lines. Fig. 9.6B diagrams the main anatomical connections extending from the retina through several stages

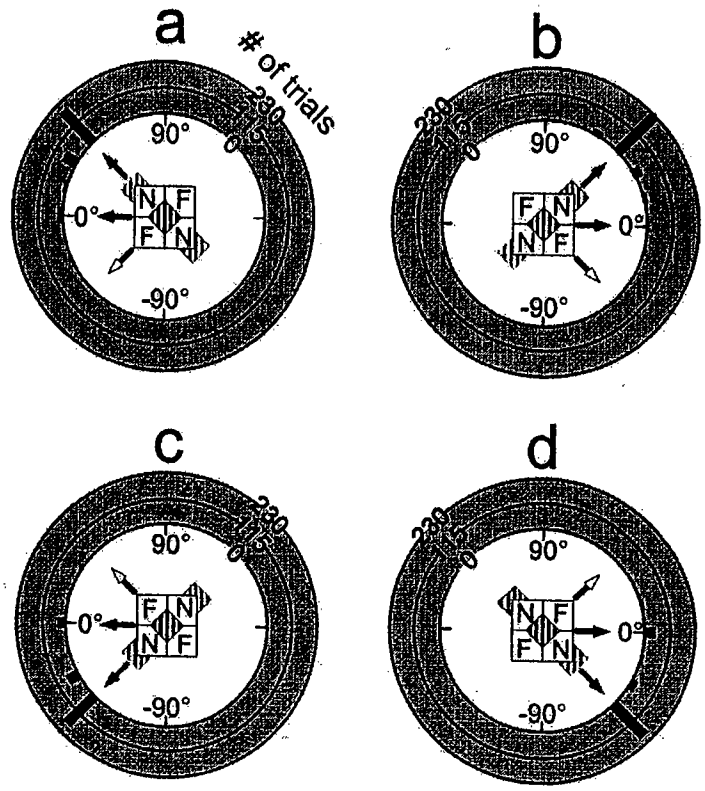

FIG. 9.5. Perceptual reports to centrally presented barber-diamonds. Stimuli were viewed within 11° apertures and possessed gratings of 0.59 cycles/deg. Data were obtained from seven human subjects. Each panel (A, B, C, and D) depicts responses to a particular barber-diamond condition as indicated by the central icons. Black arrows indicate direction of grating motion for each condition. N and F indicate crossed- and uncrossed-disparity regions that make up each depth-ordering configuration. Each barber-diamond condition consisted of a different combination of grating motion (left: A and C, right: B and D) and depthordering configuration (A and D or B and C). The implied direction of surface motion (intrinsic terminator motion) that results from these combinations is indicated by gray arrows. White arrows indicate the direction of extrinsic terminator motion. For each graph, the direction of motion reported is plotted on the polar axis, and the frequency of responses for each direction is plotted on the radial axis, (black bars). Left/right perceived motion is indicated on the horizontal axis, and up/down motion is represented on the vertical axis. Each subject participated in 40 trials per condition ($N = 1120$). Most of the reports for each condition were biased in the direction of the intrinsic terminators. From "Occlusion and the Interpretation of Visual Motion: Perceptual and Neuronal Effects of Context," by R. O. Duncan, T. D. Albright, and G. R. Stoner, 2000, *Journal of Neuroscience, 20,* p. 5890. Copyright 2000 by The Society for Neuroscience. Adapted with permission.

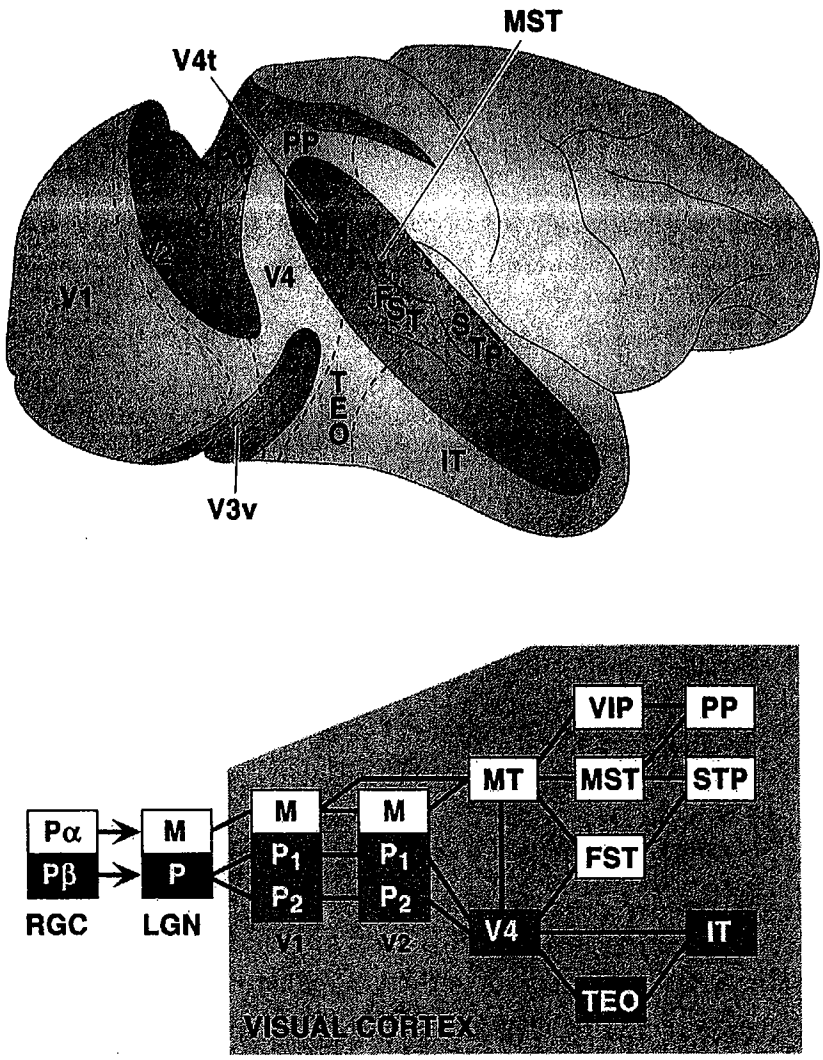

FIG. 9.6. (A) Lateral view of macaque brain showing location of striate cortex (V1) and some extrastriate visual areas. Sulci have been partially opened (shaded regions). Indicated borders of visual areas (dashed lines) are approximate. (EC, external calcarine sulcus; IO, inferior occipital sulcus; IP, intraparietal sulcus; LA, lateral sulcus; LU, lunate sulcus; PO, parieto-occipital sulcus; ST, superior temporal sulcus.) (B) Diagram of anatomical connectivity, emphasizing hierarchical organization and parallel processing streams along the geniculostriate-extrastriate pathway. Except where indicated by arrows, connections are known to be reciprocal. Not all known components of magnocellular (unshaded) and parvocellular (shaded) pathways are shown. (RGC, retinal ganglion cell layer; LGN, lateral geniculate nucleus of the thalamus; M, magnocellular subdivisions; P_1 & P_2, parvocellular subdivisions; MT, middle temporal; MST, medial superior temporal; FST, fundus superior temporal; PP, posterior parietal cortex; VIP, ventral intraparietal; STP, superior temporal polysensory). From "Cortical Processing of Visual Motion," by T. D. Albright, 1993, In: J Wallman and FA Miles (eds.) *Visual Motion and Its Role in the Stabilization of Gaze. 5,* p. 178. Copyright 1993 by Elsevier Science. Reprinted with permission.

of cortical visual processing. As previously noted, many neurons in V1 exhibit selectivity for the direction of stimulus motion (Albright, 1984; Hubel & Wiesel, 1968), but this property is particularly prominent in the middle temporal visual area (Albright, 1984), which is commonly known as area MT. Area MT is a relatively small visual area in the lower bank of the superior temporal sulcus and receives direct input from V1 (Gattass & Gross, 1981; Ungerleider & Mishkin, 1979; Van Essen, Maunsell, & Bixby, 1981). MT is thus at an early level of processing in the cortical visual hierarchy. Approximately 90% of MT neurons are selective for the direction of stimulus motion (Albright, 1984; Allman & Kaas, 1971; Maunsell & Van Essen, 1983; Zeki, 1974).

Directional responses from one such neuron are shown in Fig. 9.7. On the basis of this striking selectivity, as well as related findings (e.g., Albright, 1992; Albright,

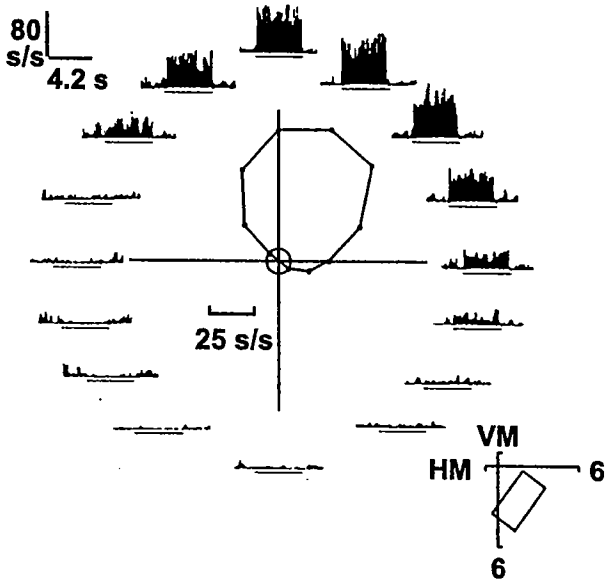

FIG. 9.7. Example of direction tuning of a typical MT neuron. Individual histograms represent responses summed over five trials to each of 16 directions of motion of a random-dot pattern moving at 20°/s. The line beneath each histogram indicates period of time during which stimulus was moving through the receptive field (shown at lower right). In the center, response to each direction is plotted on a polar graph. The polar axis represents direction of stimulus motion, the radial axis represents response (measured as spikes per second), and the small circle represents the level of spontaneous activity. The marked suppression of activity seen when the stimulus moved 180° from the optimal direction is characteristic of many MT neurons. VM, vertical meridian; HM, horizontal meridian. From "Direction and Orientation Selectivity of Neurons in Visual Area MT of the Macaque," by T. D. Albright, 1984, *Journal of Neurophysiology, 52,* p. 1109. Copyright 1984 by The American Physiological Society. Reprinted with permission.

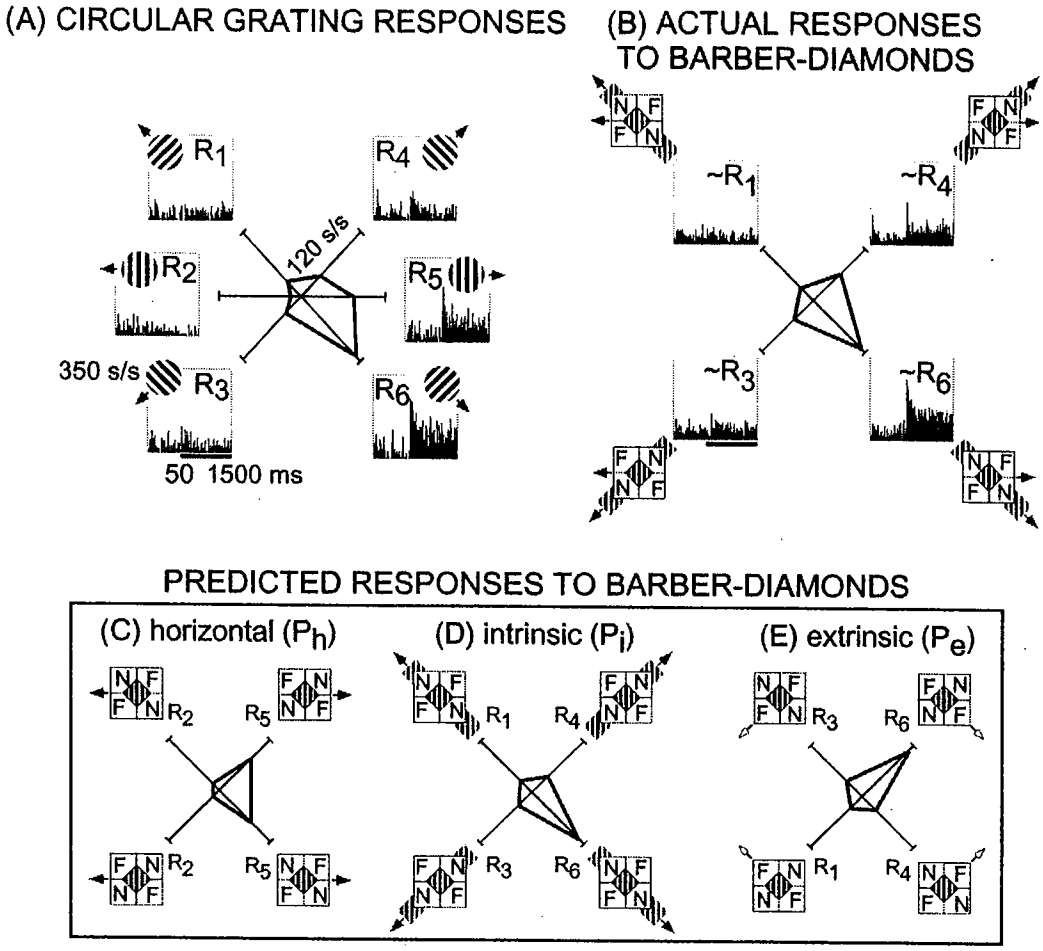

FIG. 9.8. Influence of depth ordering on direction selectivity of an MT neuron. The receptive field of this neuron was located 4° eccentric to the center of gaze and was 4.5° in diameter. (A) Peristimulus time histograms (PSTHs) illustrate neuronal responses to circular gratings moved in each of six directions. Average responses are indicated on the polar plot at center, in which polar angle corresponds to direction of stimulus motion and radial amplitude corresponds to response in spikes/second. These responses (*R*1–*R*6) were used to form three different predictions for barber-diamond stimuli (C, D, and E). (B) Actual responses to barber-diamond stimuli. PSTHs for each of the four barber-diamond conditions are presented. The average responses to moving barber-diamonds are plotted in polar coordinates at center. The bars under the lower left histograms in A and B indicate the period of stimulus movement (C) Horizontal motion prediction (P_h). Icons illustrate the stimulus configuration for each of four experimental conditions. Predicted neuronal responses to each condition are shown on the polar plot at center. This prediction holds that neuronal responses are influenced solely by the direction of grating motion (black arrows) and, hence, are of the same relative magnitude as responses to circular gratings moved leftward (R2) and rightward (R5), regardless of depth-ordering configuration. (D) Intrinsic motion prediction (P_i). This prediction holds that responses are associated with the direction of

Desimone, & Gross, 1984; Dobkins & Albright, 1994; Movshon, Adelson, Gizzi, & Newsome, 1985; Newsome, Britten, & Movshon, 1989; Newsome & Pare, 1988; Rodman & Albright, 1987; Rodman & Albright, 1989; Stoner & Albright, 1992), area MT is widely regarded as a primary component of the neural substrate for visual motion perception. For this reason, we began our investigation of the neuronal bases of contextual influences on motion perception by recording activity from isolated neurons in area MT while rhesus monkeys viewed barber-diamond stimuli (Duncan et al., 2000).

Stimuli were configured precisely as they had been for the psychophysical experiments previously described and were positioned so that the vertical bars covered the CRF of each neuron studied. We predicted that MT neurons would exhibit selectivity for the direction in which the intrinsic terminators moved, consistent with perception. An illustration of this prediction for one MT neuron is shown in Fig. 9.8. Fig. 9.8A illustrates the direction tuning of the neuron, assessed using a drifting pattern of stripes (grating) presented within a circular aperture over the classical receptive field. Average responses for gratings moved in each of six different directions and are plotted in polar coordinates at the center of Fig. 9.8A, with angle corresponding to the direction of stimulus motion. Peristimulus histograms around the perimeter of the plot show summed responses as a function of time. This cell responded most strongly to motion downward and rightward.

From the data in Fig. 9.8A, we considered three different predictions for the responses of this neuron to barber-diamonds. For each prediction (illustrated in Fig. 9.8C, 9.8D, and 9.8E), the expected responses to the four barber-diamond conditions are shown. The simplest prediction, shown in Fig. 9.8C, is that of the null hypothesis, which assumes that the depth configuration (i.e., the context) has no influence over the directional selectivity of the cell. If true, the neuronal response would be determined exclusively by the leftward or rightward direction of motion of the barber-diamond stripes. The predicted responses to both leftward

FIG. 9.8. (*continued*) intrinsic terminator motion (gray arrows), and hence are of the same relative magnitude as responses to circular grating moving in the corresponding oblique directions (R1, R3, R4, and R6). (E) Extrinsic motion prediction (P_e). This prediction holds that responses will be associated with the direction of extrinsic terminator motion (white arrows), and hence be of the same relative magnitude as the intrinsic motion prediction but reflected about the horizontal axis. Observed responses (panel B) of this neuron to barber-diamond stimuli ($R_{i|h} = 0.85$) were more closely correlated with the intrinsic motion prediction than with either the horizontal or extrinsic motion prediction. From "Occlusion and the Interpretation of Visual Motion: Perceptual and Neuronal Effects of Context," by R. O. Duncan, T. D. Albright, and G. R. Stoner, 2000, *Journal of Neuroscience, 20,* p. 5889. Copyright 2000 by The Society for Neuroscience. Adapted with permission.

motion conditions are thus equal to the responses to the corresponding leftward motion (*R*2) seen in Fig. 9.8A. Similarly, the predicted responses to rightward motion conditions are equal to *R*5.

The intrinsic motion prediction is shown in Fig. 9.8D. According to this prediction, neuronal responses to the four barber-diamond conditions reflect the four directions of intrinsic terminator motion and, hence, will match the pattern of responses seen for the gratings moving in those four directions. Specifically, the four predicted responses are equal to the responses measured for the corresponding directions (*R*1, *R*3, *R*4, and *R*6) seen in Fig. 9.8A. To cover all possibilities, we also considered the extrinsic motion prediction shown in Fig. 9.8E. Because the extrinsic terminators always move orthogonally to the intrinsic terminators in the barber-diamond, the extrinsic prediction is simply the intrinsic prediction reflected about the horizontal axis.

The responses of this neuron to the barber-diamond stimuli are shown in Fig. 9.8B. The pattern of selectivity clearly violated the null hypothesis because the responses were not independent of the depth-ordering configuration (e.g., the response to the upper right barber-diamond condition was much smaller than that elicited by the lower right stimulus). Moreover, as predicted, the largest response occurred when the intrinsic terminators moved in the preferred direction for the cell (downward and rightward, see Fig. 9.8A). Indeed, the shape of the direction tuning curve obtained with barber-diamond stimulation more closely matched that of the intrinsic prediction (Fig. 9.8D), than it did either the horizontal (Figure 9.8C) or extrinsic (Figure 9.8E) prediction. To quantify the accuracy of our predictions, we computed a measure of the correlation between the intrinsic prediction and the actual barber-diamond data ($R_{i|h}$).[1] $R_{i|h}$ varies between –1 and +1, with larger positive values reflecting increased correlation with the intrinsic predictor. The value of the coefficient $R_{i|h}$ for the cell documented in Fig. 9.8 is 0.85.

Several additional examples of the effect of barber-diamond stimulation are shown in Fig. 9.9. For five of the six cases shown, responses to the barber-diamonds (gray lines) were positively correlated with the intrinsic predictor (black lines) (values of $R_{i|h}$ range from 0.54 to 0.87). The sixth case (lower right) illustrates that, although it may be common, this phenomenon is not universal: Here the barber-diamond response was negatively correlated with the intrinsic prediction (and thus positively correlated with the extrinsic prediction).

Fig. 9.10 shows the distribution of correlation coefficients ($R_{i|h}$). Plotted are the values of $R_{i|h}$ for all of the MT neurons studied ($n = 265$; white bars), as well as for the subset of sampled neurons for which the values of $R_{i|h}$ were significantly different from 0 ($n = 90$; gray bars). The distribution as a whole is shifted toward positive coefficients, with the average $R_{i|h}$ (0.10) being significantly different

[1]Because the horizontal motion prediction is itself correlated with the intrinsic prediction, we computed a partial correlation coefficient for the intrinsic predictor with the horizontal predictor partialed out. Conveniently, the partial correlation coefficient for the extrinsic predictor is equal in magnitude to that for the intrinsic predictor but of opposite sign. See Duncan et al. (2000) for details.

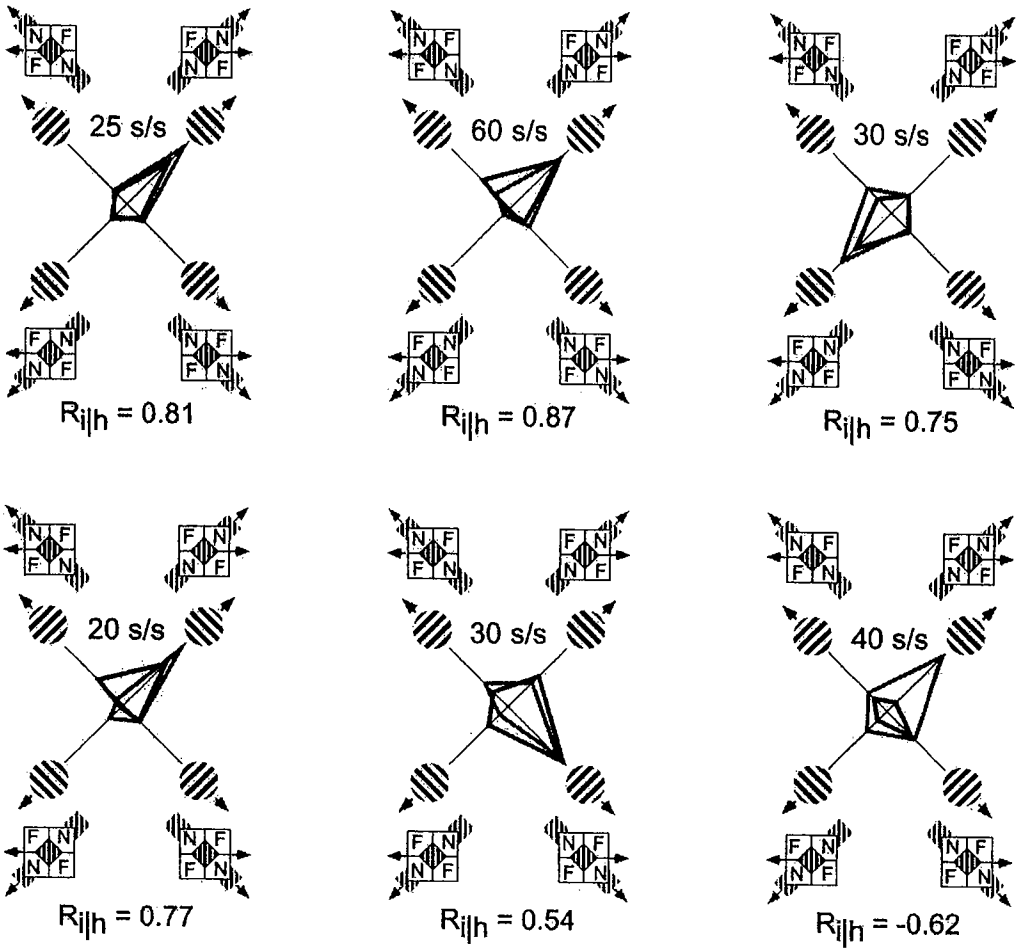

FIG. 9.9. Barber-diamond data from six MT neurons, each of which demonstrated significant influences of depth ordering on directional responses. Five of these neurons had positive intrinsic correlation coefficients and one neuron had a negative coefficient (*lower right*). Responses to circular gratings were averaged across five trials and plotted in polar coordinates (black lines). The direction of motion for each grating condition is indicated by the icons along the polar axis and the mean response to each condition is plotted along the radial axis. Responses to each barber-diamond condition were averaged across 10 trials and plotted along with corresponding icons on the same graphs (gray lines). Each of these neurons demonstrated individually significant responses to barber-diamonds that were configured to elicit upward versus downward motion (ANOVA; all $p < 0.0004$). Neurons with positive coefficients ($R_{i|h}$) have directionally selective responses consistent with the direction of motion of the intrinsic terminators. See also caption to Fig. 9.8. From "Occlusion and the Interpretation of Visual Motion: Perceptual and Neuronal Effects of Context," by R. O. Duncan, T. D. Albright, and G. R. Stoner, 2000 *Journal of Neuroscience, 20,* p. 5891. Copyright 2000, by The Society for Neuroscience. Adapted with permission.

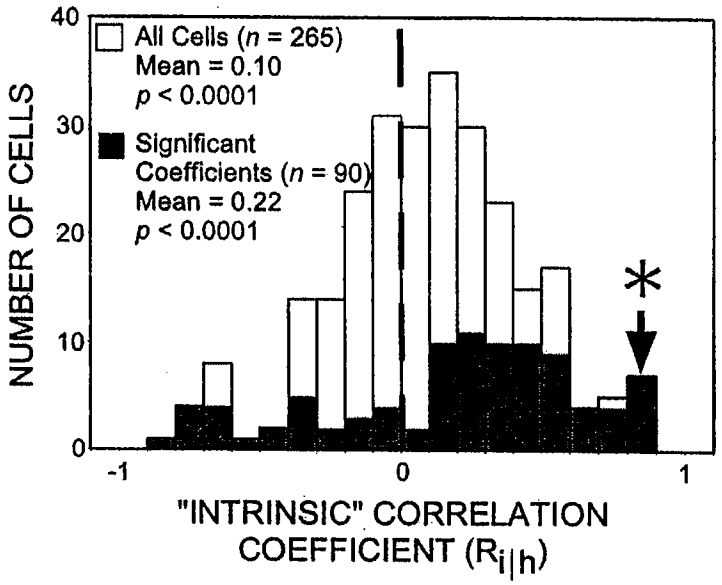

FIG. 9.10. Neuronal population data from barber-diamond studies. Distribution of intrinsic correlation coefficients ($R_{i|h}$) for the sample of 265 MT neurons studied (unfilled bars). The subset of cells for which coefficients reached statistical significance (ANOVA; all $p < 0.05$) is indicated by the gray bars. Positive coefficients reflect a positive correlation with the motion of intrinsic terminators. The population displays a significant positive shift of the mean (0.10) for intrinsic correlation coefficients (t test; $p < 0.0001$). The population of individually significant cells displays a larger number of positive coefficients ($n = 67$) relative to negative ones ($n = 23$). Asterisk denotes coefficient for neuron illustrated in Fig. 9.7. For the entire sample of 265 neurons, mean eccentricity of receptive field centers was 5.5°, and mean receptive field diameter was 5.4°. From "Occlusion and the Interpretation of Visual Motion: Perceptual and Neuronal Effects of Context," by R. O. Duncan, T. D. Albright, and G. R. Stoner, 2000, *Journal of Neuroscience, 20,* p. 5892. Copyright 2000 by The Society for Neuroscience. Adapted with permission.

from 0. Moreover, the distribution of individually significant coefficients is skewed toward positive values (mean $R_{i|h} = 0.22$). Thus, for at least a subset of MT neurons, we observed a high degree of correspondence between the intrinsic predictor and barber-diamond responses (Duncan et al., 2000), supporting our hypothesis that context and feature interpretation influence the signals produced by cortical neurons (Stoner & Albright, 1993). More generally, these data provide striking evidence that some MT neurons encode the perceived motions of surfaces, rather than indiscriminately representing the motions of isolated retinal image features.

It is reassuring to find that the signals of neurons at relatively early stages of visual processing represent the structure of the world—the things perceived—as opposed to undifferentiated retinal image features. These findings encourage us

to search for the mechanisms by which the representational transformation comes about. This will be a major focus of research in coming years.

Feature Grouping

The experiments previously described addressed how motion processing is influenced by the assignment of features to particular objects in a scene. We now turn to a related problem: how motion processing is influenced by the perceptual grouping of image features that are visually related to each other—for example, the grouping of features that have the same color.

To appreciate how grouping influences motion processing, consider the following real-world problem. You are standing on the balcony of Grand Central Station during a busy rush hour looking down on the concourse, upon which hundreds (if not thousands) of people are moving in many different directions. Through this crowd moves a scattered group of schoolchildren, and you are hoping to identify the exit for which they are heading so that you can meet them there. Under normal circumstances, owing to the density of people, their lack of individual distinctiveness, and the complexity of their motions, it would be difficult to track the motion of the schoolchildren. The schoolchildren, however, wear distinctive red hats, making them easy to detect. Moreover, you find that even in the presence of abundant dynamic visual noise (the other travelers, slight direction changes by the individual children as they navigate through the crowd), it is a simple matter to determine the average direction of the children as they move through the crowd. Intuitively, it seems that the chromatic distinctiveness of signal and noise elements in the scene enables them to be grouped independently and facilitates the processing of their motion. Moreover, context is a critical parameter. In this case, the relevant context is the color of the signal elements relative to that of the noise.

Although this contextual effect on motion processing is easy to understand intuitively, it poses a problem for visual neurophysiology because for many years it has been thought that motion-sensitive neurons are blind to chromatic information. Earlier experiments on color selectivity and motion processing have not, however, manipulated the chromatic context in which stimuli are moved: hence, they are irrelevant to the phenomenon exemplified in the real-world example. To elucidate the neuronal bases of the contextual effect of color on motion processing, we therefore performed a series of studies that included psychophysical experiments with human subjects and paired neurophysiological/psychophysical experiments with rhesus monkeys.

The Stochastic Motion Stimulus

For these experiments we used a visual stimulus that both captures essential features of the real-world problem previously described and can be manipulated parametrically. This stimulus (Fig. 9.11) is based on one used in previous studies

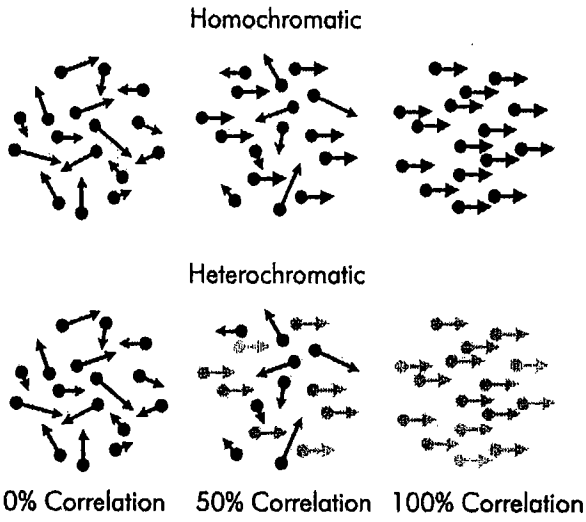

FIG. 9.11. Schematic depiction of the stimulus used to study the influence of chromatic context on perceived motion and its neural correlates. Each stimulus consisted of a sequence of frames of randomly positioned bright dots appearing against a dark background on a CRT screen. Dots in each of the six circular apertures of the figure represent dots in six different stimuli. Arrows indicate velocity (direction and speed). The proportion of dots moving in the same direction at the same speed, expressed as a percentage and termed the motion coherence describes the strength of the motion signal. At 0% motion coherence, all of the dots move randomly. At 50% motion coherence, half the dots have the same velocity. At 100% motion coherence, all of the dots have the same velocity. In the homochromatic condition, all of the dots have the same color. In the heterochromatic condition, the signal dots are a different color from the noise dots. From "Segmentation by Color Influences Responses of Motion-Sensitive Neurons in the Cortical Middle Temporal Visual Area," by L. J. Croner and T. D. Albright, 1999, *Journal of Neuroscience, 19,* p. 3936. Copyright 1999 by The Society for Neuroscience. Reprinted with permission.

by Newsome and colleagues (e.g., Newsome et al., 1989; Newsome & Pare, 1988; Salzman, Britten, & Newsome, 1990), designed to quantify the motion sensitivity of primate observers and cortical neurons. Details of stimulus construction and display are provided in Croner and Albright (1997, 1999a). Briefly, the stimulus (Fig. 9.11A) used in previous studies consisted of a patch of randomly positioned dots that were displaced on each temporal frame. A variable fraction of these dots—the signal—as displaced by a common motion vector. The remaining dots—the noise—were displaced randomly and with uncorrelated motion vectors. Motion signal strength is the proportion of correlated motions in the stimulus and could be varied continuously from 0% (the motion of each dot is independent of the motions of all other dots) to 100% (the motion of all dots is in the same direction

and speed). Because signal and noise dots were otherwise identical (i.e., the same color), we refer to this traditional stochastic motion stimulus as homochromatic. Using this stimulus, Newsome and colleagues showed that the ability of human and nonhuman primate observers to discriminate direction of motion exhibits a reliable dependence on motion signal strength (Newsome & Pare, 1988). In addition, these investigators demonstrated that the directionally selective responses of individual MT neurons exhibit a similar dependence on motion signal strength (Newsome et al., 1989).

Our adaptation of the traditional stochastic motion stimulus was composed of signal dots of one color (e.g., red) and noise dots of a different color (e.g., green) (Fig. 9.11B). This heterochromatic version is identical to the homochromatic version in all respects, save the chromatic difference between signal and noise dots.

Feature Grouping by Color Improves Perceptual Discrimination of Motion in Visual Noise

The analogy between our heterochromatic stimulus and the real-world example previously described needs little explication. Tracking the global motion of the schoolchildren was simplified because of their red hats. Similarly, we predicted that it would be easier to discern the direction of signal motion in the heterochromatic condition relative to the homochromatic condition. Specifically, we predicted that for a given motion signal strength, direction discriminability would improve (provided it was not already at ceiling), and the discriminability threshold for heterochromatic stimuli would be markedly lower than for homochromatic stimuli (Croner & Albright, 1997).

Five human subjects viewed both homo- and heterochromatic motion stimuli (Fig. 9.11), which were randomly interleaved from trial to trial. Both stimulus types were presented at a number of different motion signal strengths, spanning the extremes of behavioral performance including threshold. Color (red vs. green) assignments for signal and noise dots in heterochromatic stimuli were randomized across trials. Subjects fixated a small target while each stimulus was presented at the center of gaze for 2s. Signal motion was either leftward or rightward on each trial, following a pseudorandom schedule. At the conclusion of each presentation, subjects indicated the perceived direction of signal motion (2-Alternative Forced-Choice; left vs. right).

Data from one human subject are shown in Fig. 9.12. Task performance (proportion of correct direction judgments) is plotted as a function of motion signal strength for homochromatic (open triangles) and heterochromatic (filled circles) stimuli. Consistent with previous findings (Newsome et al., 1989), directional discriminability increased with motion signal strength, as revealed by the sigmoidal psychometric functions. The effect of the heterochromatic manipulation manifested as a pronounced leftward shift of the psychometric function. This reflects

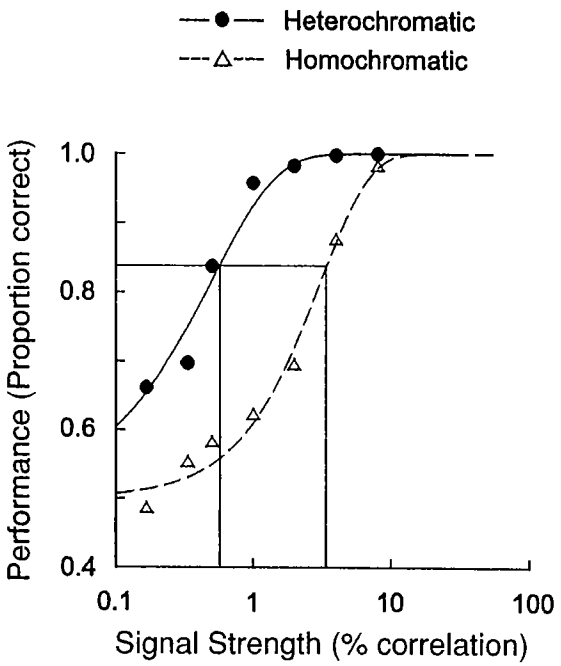

FIG. 9.12. Performance functions measured for a representative human subject discriminating signal direction in homochromatic (open triangles) and heterochromatic (filled circles) stimuli. Also shown are curves fit to the homochromatic (dashed lines), and heterochromatic (solid lines) data. Threshold performance (0.82) is illustrated by a thin horizontal line. Where this line intersects each psychometric function, a thin vertical line is drawn to intersect the *x*-axis at the threshold motion coherence for the function. Threshold motion coherence for the homochromatic condition was 3.1%, and for the heterochromatic condition it was 0.5%, indicating an approximately sixfold decrease in threshold when signal and noise dots were distinguished by color. From "Image Segmentation Enhances Discrimination of Motion of Visual Noise," by L. J. Croner and T. D. Albright, 1997, *Vision Research, 37*, p. 1418. Copyright 1997 by Elsevier Science. Adapted with permission.

the predicted improved sensitivity to motion direction in hetero- versus homochromatic stimuli: The discriminability threshold was significantly reduced for hetero- (0.5%) versus homochromatic stimuli (3.1%). Although the magnitude of threshold reduction varied across human subjects, significant effects of chromatic context on thresholds for direction discrimination were observed in every case.

From these psychophysical findings we speculated that the effects of grouping by color cues on motion perception were implemented by a dynamic gating mechanism. Specifically, we hypothesized that, during each stimulus presentation, the visual system created a chromatic filter that blocked information about noise dots from reaching motion detection circuits in the visual cortex (Croner & Albright, 1997). This simple hypothesis, of course, begs important questions regarding the

mechanism by which such a filter may be implemented. Answers to these questions must await additional research.

Feature Grouping by Color Improves Neuronal Discrimination of Motion in Visual Noise

Using the homochromatic version of the stochastic motion stimulus, Newsome and colleagues found that the discriminability of direction by individual MT neurons closely matched monkeys' performance on the direction discrimination task (Newsome et al., 1989). These investigators thus hypothesized that the monkeys' perceptual decisions were based on the signals provided by MT neurons. If this hypothesis were correct, we would expect that stimulus manipulations influencing psychophysical performance would also influence neuronal discriminability. Because we found that chromatic context influences direction discrimination performance by humans, it follows that there may be parallel changes in the sensitivity of MT neurons. To test this, we trained monkeys to discriminate direction using both homo- and heterochromatic stochastic motion stimuli and then recorded from individual MT neurons while animals performed the task. Aside from the key stimulus differences, our behavioral paradigm was identical to that used by Newsome and colleagues (Newsome et al., 1989; Newsome & Pare, 1988) and is summarized schematically in Fig. 9.13.

The first goal of these experiments was to determine whether chromatic context improves motion discrimination thresholds in monkeys as it does in humans. Fig. 9.14 illustrates that not only did monkeys exhibit a consistent decrease in psychometric threshold for hetero- versus homochromatic stimuli, but also the effect generalized across a range of visual field eccentricities and axes of motion. This generalization was critical because it subsequently enabled us to customize the stochastic motion stimuli to the receptive field position and direction preference of individual MT neurons. (For a more complete comparison of behavioral data from monkeys and humans, see Croner & Albright, 1999a.)

We next examined the effects of these manipulations on the directional selectivity of individual MT neurons (Croner & Albright, 1999a). Stimuli were centered on the CRF of each MT neuron studied, and the two directions of signal motion were aligned with the neuron's preferred and antipreferred directions. Color (red vs. green) assignments for signal and noise dots in heterochromatic stimuli were randomized across trials. To quantify each neuron's ability to discriminate motion direction, we adopted the methods of signal detection theory (Swets, Green, Getty, & Swets, 1978) previously used for similar purposes by Newsome and colleagues (Newsome et al., 1989). Briefly, we examined the distributions of responses elicited by preferred and antipreferred directions of each stimulus condition (each signal strength of homo- or heterochromatic stimuli) and performed an analysis of receiver operating characteristics (ROC) based on these responses. We then computed the probability that an ideal observer could correctly determine the direction

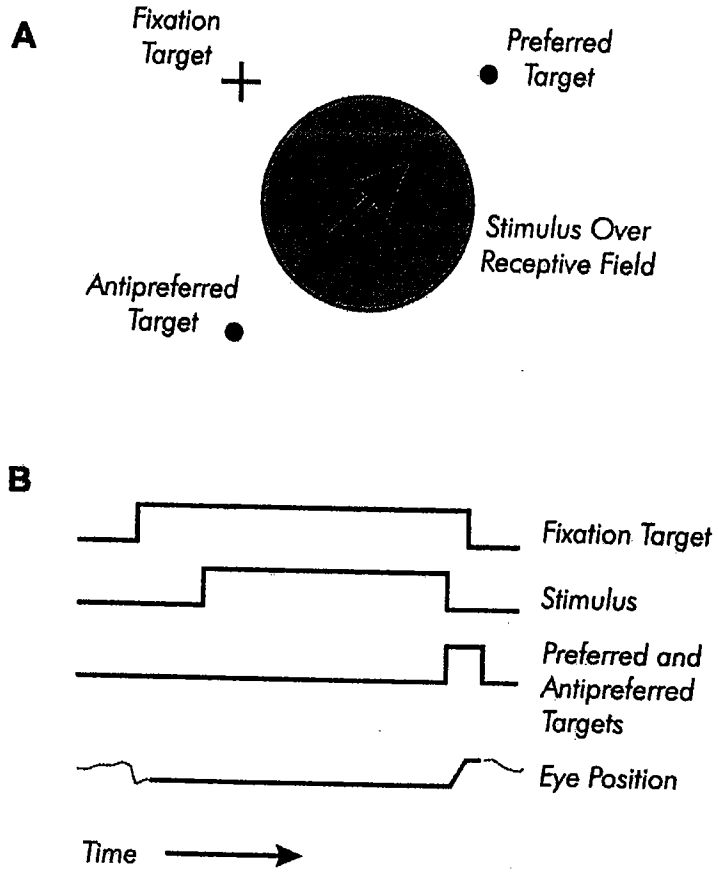

FIG. 9.13. Schematic depiction of paradigm developed by Newsome and colleagues (e.g., Newsome et al., 1989) and subsequently used by Croner and Albright (1999) to investigate influence of feature grouping on motion processing. (A) Example spatial configuration of the fixation target, stimulus aperture, and targets for direction choice. The stimulus aperture was superimposed on the receptive field of the neuron under study. Signal motion during each trial was in either the neuron's preferred or antipreferred direction; the targets for direction choice were positioned according to the neuron's preferred direction. (B) Diagram of the temporal sequence of events during one trial. The trial was initiated with the onset of the fixation target (Fixation Target). Five hundred ms after fixation was achieved (Eye Position), the motion stimulus was presented for 2 s (Stimulus). When the stimulus was extinguished, the Preferred and Antipreferred Targets appeared and remained on until the monkey indicated its direction choice by making a saccadic eye movement to one of them. From "Segmentation by Color Influences Responses of Motion-Sensitive Neurons in the Cortical Middle Temporal Visual Area," by L. J. Croner and T. D. Albright, 1999, *Journal of Neuroscience, 19,* p. 3938. Copyright 1999 by The Society for Neuroscience. Reprinted with permission.

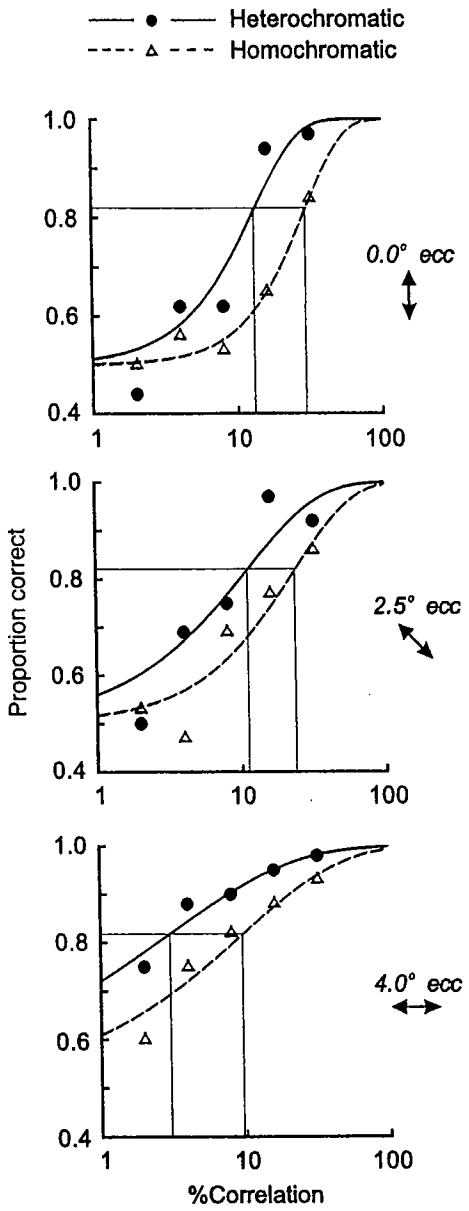

FIG. 9.14. Performance data obtained from monkeys discriminating signal direction in homochromatic (open triangles) and heterochromatic (filled circles) stimuli. Also shown are curves fit to the homochromatic (dashed lines) and heterochromatic (solid lines) data. The inset to the right of each plot gives the retinal eccentricity of the stimulus and the axis of signal direction used in each block. Thin straight lines indicate performance thresholds, as in Fig. 9.12. The homochromatic and heterochromatic psychophysical thresholds, respectively, were 29.7% and 13.2% (top), 23.5% and 10.9% (middle), and 9.5% and 2.9% (bottom). From "Segmentation by Color Influences Responses of Motion-Sensitive Neurons in the Cortical Middle Temporal Visual Area," by L. J. Croner and T. D. Albright, 1999, *Journal of Neuroscience, 19,* p. 3939. Copyright 1999 by The Society for Neuroscience. Reprinted with permission.

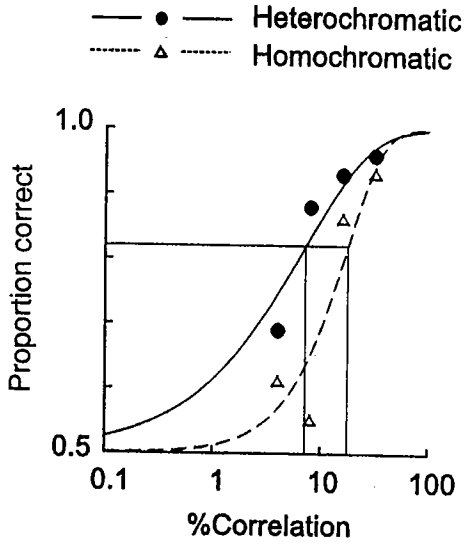

FIG. 9.15. Paired neurometric functions obtained from one neuron's responses to the homo- and heterochromatic stimuli. In this experiment, we measured significantly different neuronal thresholds for the two conditions. Thin straight lines illustrate thresholds, as in Fig. 9.12. The homochromatic and heterochromatic neurometric thresholds were, respectively, 17.6 and 7.0%. From "Segmentation by Color Influences Responses of Motion-Sensitive Neurons in the Cortical Middle Temporal Visual Area," by L. J. Croner and T. D. Albright, 1999, *Journal of Neuroscience, 19,* p. 3942. Copyright 1999 by The Society for Neuroscience. Reprinted with permission.

of stimulus motion given the observed responses for each condition. These probability estimates served as an index of neuronal directional discriminability and were used to generate a neurometric function relating discriminability to motion signal strength for each cell. Fig. 9.15 shows the neurometric functions for homo- (open triangles) and heterochromatic stimuli (filled circles) obtained from one MT neuron. This neuron exhibited a dramatic threshold reduction for heterochromatic stimuli, paralleling the typical psychophysical threshold reduction.

To convey the distribution of this effect on neuronal discriminability, we include Fig. 9.16, which contains a plot of homo- versus heterochromatic threshold for each neuron studied. Although the thresholds were significantly different—like the example in Fig. 9.15—in only a subset of cases (filled circles), the average ratio of hetero- to homochromatic threshold (0.74) was significantly less than 1.0 (see diagonal plot), and the ratio for the subset of individually significant neurons (0.35, black bars) revealed a strong effect of chromatic context.

Relative to the homochromatic stimulus, the heterochromatic stimulus thus elicited lower average psychophysical and neuronal thresholds for motion discriminability. These findings support the hypotheses that (1) MT neurons account for the perceptual discriminability of motion direction with chromatic grouping

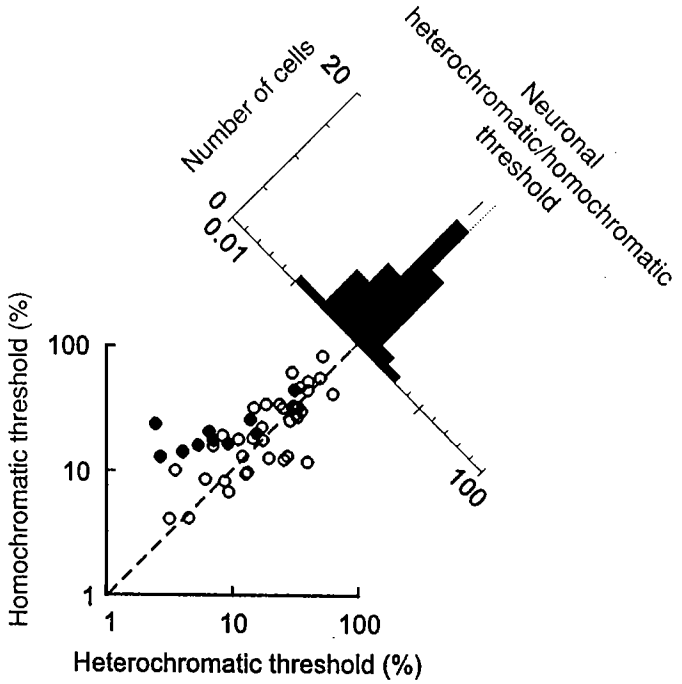

FIG. 9.16. Comparison of neuronal performance for homochromatic and heterochromatic conditions. The bottom panel shows a scatterplot of the absolute thresholds obtained in experiments with single neurons. The black symbols signify neurons for which the two thresholds were significantly different from each other; the broken line illustrates where points would fall if the thresholds were identical. The top right panel shows a frequency distribution of the ratios of heterochromatic to homochromatic thresholds, formed by summing across the scatterplot within diagonally oriented bins. The dotted line indicates unity, and the solid line segment is aligned with the geometric mean. Ratios less than unity indicate that neuronal performance was better (threshold was lower) for the heterochromatic condition. The black bars show the threshold ratios for experiments in which the two thresholds were significantly different from each other. From "Segmentation by Color Influences Responses of Motion-Sensitive Neurons in the Cortical Middle Temporal Visual Area," by L. J. Croner and T. D. Albright, 1999, *Journal of Neuroscience, 19,* p. 3943. Copyright 1999 by The Society for Neuroscience. Reprinted with permission.

of signal versus noise and (2) these neurons have access to contextual (chromatic, in this case) information unrelated to motion per se. Because both psychophysical and neuronal data were obtained on each trial, however, we can address a still finer question: Do fluctuations in psychophysical performance correlate with fluctuations in neuronal discriminability? We approached this question by assessing the change in psychophysical threshold as a function of the change in the neuronal threshold for each block of trials. The results are plotted in Fig. 9.17. Here, each data set represents a complete experiment for one neuron, with the four relevant

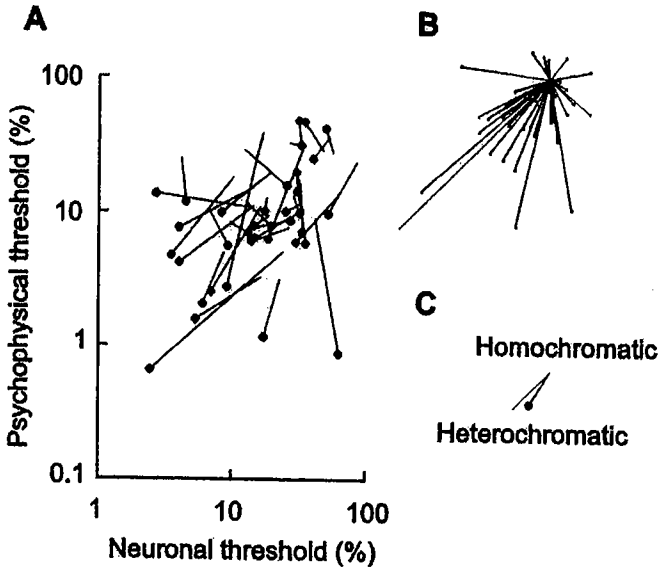

FIG. 9.17. Comparison of the change in absolute neuronal and behavioral thresholds afforded by color segmentation. (A) Vectors show the change in thresholds measured in each experiment. The plain end of each vector shows the relation between behavioral and neuronal thresholds for the homochromatic condition, and the end with a black dot shows the same relation for the heterochromatic condition. Vectors with a downward component (from homochromatic to heterochromatic) indicate enhanced behavioral sensitivity to the heterochromatic condition; vectors with an upward component indicate the converse. Vectors with a leftward component indicate enhanced neuronal sensitivity to the heterochromatic condition; vectors with a rightward component indicate the converse. (B) The vectors are redrawn from the same origin, which represents the homochromatic thresholds. (C) The single vector is the average of the vectors shown in B. In B and C the dotted line is the 45° diagonal, where vectors would lie if color segmentation influenced behavioral and neuronal thresholds equally. From "Segmentation by Color Influnces Responses of Motion-Sensitive Neurons in the Cortical Middle Temporal Visual Area," by L. J. Croner and T. D. Albright, 1999, *Journal of Neuroscience, 19,* p. 3944. Copyright 1999 by The Society for Neuroscience. Reprinted with permission.

thresholds (psycho-hetero, psycho-homo, neuro-hetero, neuro-homo) represented by a single vector in a two-dimensional space (Fig. 9.17A). In this figure, psychophysical threshold is plotted as a function of neuronal threshold. The plain end of each vector shows the psychophysical/neuronal threshold pair obtained using the homochromatic stimulus, and the end with the black dot shows the corresponding threshold pair obtained using the heterochromatic stimulus. Vectors with vertical and horizontal components of equal direction and magnitude indicate experiments in which the stimulus manipulation had equivalent effects on the neuron and the behavior. To facilitate inspection of these data, the vectors are plotted from a

common homochromatic origin in Fig. 9.17B. Clearly, not all neurons exhibit threshold shifts that match the perceptual change. However, from this plot, one can more fully appreciate the trend, which is toward neuronal and perceptual decreases of similar magnitude (though not identical; see Croner & Albright, 1999a). The point is further emphasized in Fig. 9.17C, where we have plotted the average data vector (solid line) and the neuronal-perceptual identity vector (dashed line). On average, the introduction of a chromatic context that enables feature grouping elicits comparable changes in simultaneously assessed indices of neuronal and perceptual sensitivity.

Sources of Enhanced Directional Discriminability by Cortical Neurons

To understand the mechanisms responsible for the context-dependent changes in neuronal thresholds, we evaluated the response changes underlying the observed threshold reduction. Neuronal directional discriminability, as we measured it, depends on the degree of overlap between the distributions of responses to preferred and antipreferred directions of motion. The degree of overlap is affected by two simple attributes of these distributions: average response and variance. All else being equal, divergence of the means (by preferred direction response increase, antipreferred direction response decrease, or both) leads to greater discriminability. Similarly, reduction of the variance (for either preferred or antipreferred responses, or both) leads to greater discriminability.

Britten and colleagues (Britten, Shadlen, Newsome, & Movshon, 1992) demonstrated that, for homochromatic stimuli, improvement of neuronal discriminability with increasing signal strength results from both an increase in the magnitude of the preferred direction response and a decrease in the magnitude of the antipreferred direction response. Analysis of our homochromatic data confirmed that finding. By contrast, we found that the improved discriminability for heterochromatic stimuli, relative to that for homochromatic stimuli of the same signal strength, was associated with altered responses only to the preferred direction stimuli. Specifically, heterochromatic preferred direction stimuli evoked both larger and less variable responses than did preferred direction homochromatic stimuli of the same signal strength. These findings are summarized schematically in Fig. 9.18A. To facilitate comparison of these response changes with those elicited by increases in signal strength, we have illustrated the latter schematically in Fig. 9.18B. This comparison reveals an important point: Although either the addition of chromatic context or an increase in signal strength can improve neuronal discriminability, the underlying mechanisms are very different.

Precisely what those mechanisms are remains unclear, but the reduction of variance associated with heterochromatic stimulation is provocative, suggesting that the neuron is driven more effectively for the heterochromatic stimuli. We propose that the observed response changes reflect a unique decision strategy being used by the monkey when features are grouped by chromatic cues (or any other

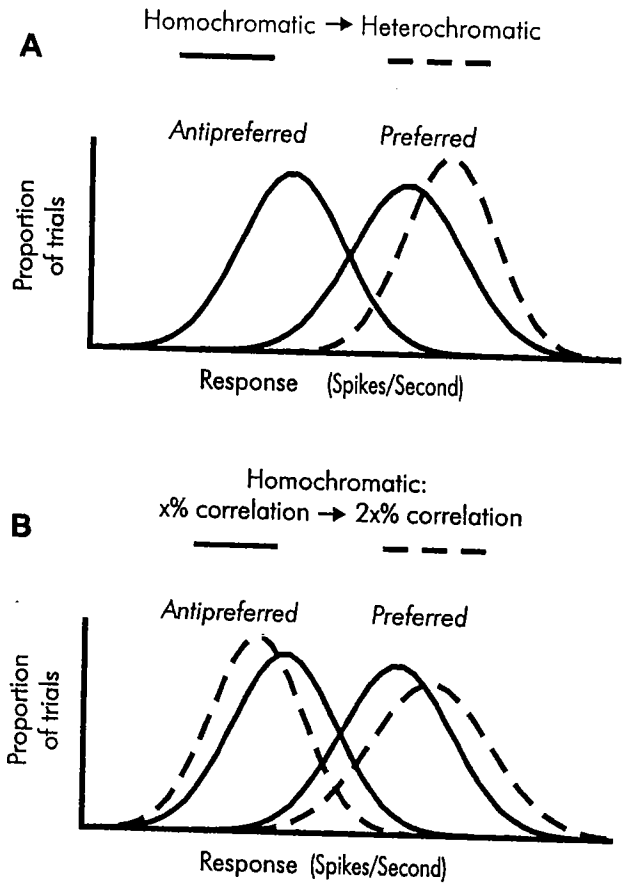

FIG. 9.18. Schematic diagrams showing the response changes associated with enhanced discriminability by neurons with significantly different homochromatic and heterochromatic thresholds. (A) Hypothetical frequency distributions of the responses of a neuron to homochromatic stimuli (solid lines) and to heterochromatic stimuli (dashed lines) of one stimulus motion coherence. The horizontal axis in each graph represents response strength (spikes/s). Each curve represents a distribution of response strengths across multiple trials, in which motion was either in the preferred or antipreferred direction. (B) Hypothetical frequency distributions of the responses of a neuron to homochromatic stimuli of a low stimulus motion coherence (solid lines) and to homochromatic stimuli of twice that value (dashed lines). From "Segmentation by Color Influences Responses of Motion-Sensitive Neurons in the Cortical Middle Temporal Visual Area," by L. J. Croner and T. D. Albright, 1999, *Journal of Neuroscience, 19,* p. 3948. Copyright 1999 by The Society for Neuroscience. Reprinted with permission.

cues allowing feature grouping). One possibility is that the chromatic cue enables attention to be directed selectively to signal dots, thus freeing the discriminative apparatus from concern with the noise dots. A precedent for this general type of mechanism can be found in the fact that attentional selection by color has a gating effect on responses in area V4 (Motter, 1994).

CONCLUSIONS

We described two sets of experiments showing that the responses of some MT neurons are influenced by the spatial context in which moving stimuli appear. In both cases, nonmotion cues (depth ordering for the first set of experiments and color for the second set) clarify the evaluation of motion signals that would be ambiguous in the domain of image motion alone. Based on these findings, we conclude that nonmotion attributes allow the visual system to select relevant image motions—that of intrinsic line terminators in the first set of experiments and of signal dots in the second set—for further processing leading to perception.

Our experiments contribute to a growing body of evidence that the responses of neurons relatively early in the visual processing hierarchy underlie perceptual organization. What remains unknown are the mechanisms by which these responses arise. Detailed mechanistic studies face formidable technical challenges (consider, e.g., the problem of revealing local anatomical circuitry in the functioning visual cortex).

Nevertheless, some progress has been made. For example, that visual context affects scene perception in particular ways implies that visual neurons have ready access to specific contextual information, and this constrains the kind of information flow we should expect to find in the visual cortex. In some cases, the process of perceptual organization seems to invoke attentional mechanisms, the neuronal bases of which can be explored by direct manipulation of the observer's attentional focus (e.g. Reynolds, Chelazzi, & Desimone, 1999).

Regardless of the exact mechanisms involved, the insight that contextual information has specific influences over neuronal representation in the visual cortex significantly affects our understanding of how the visual system is organized. In addition, this insight signals that an important bridge has been crossed: Neurobiologists and psychologists have begun to speak in the same terms, value the same concepts, and target the same problems of perception. It is only a matter of time before we truly understand how the brain accomplishes visual perception.

ACKNOWLEDGMENTS

We thank J. Costanza for superb technical assistance and Carl Olson and Rudiger von der Heydt for comments on the manuscript. The work reviewed herein was supported, in part, by grants from the NEI (TDA, GRS), the NIMH (TDA), an

individual NRSA (LJC), an award from the Fight for Sight research division of Prevent Blindness America (LJC), and a fellowship from the McDonnell-Pew Center for Cognitive Neuroscience at San Diego (GRS). T. D. Albright is an investigator of the Howard Hughes Medical Institute.

REFERENCES

Albright, T. D. (1984). Direction and orientation selectivity of neurons in visual area MT of the macaque. *Journal of Neurophysiology, 52*(6), 1106–1130.

Albright, T. D. (1992). Form-cue invariant motion processing in primate visual cortex. *Science, 255*(5048), 1141–1143.

Albright, T. D. (1993). Cortical processing of visual motion. *Reviews of Oculomotor Research, 5,* 177–201.

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

Allman, J. M., & Kaas, J. H. (1971). Representation of the visual field in striate and adjoining cortex of the owl monkey (*Aotus Trivirgatus*). *Brain Resarch., 35,* 89–106.

Britten, K. H., Shadlen, M. N., Newsome, W. T., & Movshon, J. A. (1992). The analysis of visual motion: a comparison of neuronal and psychophysical performance. *Journal of Neuroscience, 12*(12), 4745–1465.

Croner, L. J., & Albright, T. D. (1997). Image segmentation enhances discrimination of motion in visual noise. *Vision Research, 37*(11), 1415–1427.

Croner, L. J., & Albright, T. D. (1999a). Segmentation by color influences responses of motion-sensitive neurons in the cortical middle temporal visual area. *Journal of Neuroscience, 19*(10), 3935–3951.

Croner, L. J., & Albright, T. D. (1999b). Seeing the big picture: integration of image cues in the primate visual system. *Neuron, 24*(4), 777–789.

Dobkins, K. R., & Albright, T. D. (1994). What happens if it changes color when it moves? The nature of chromatic input to macaque visual area MT. *Journal of Neuroscience, 14*(8), 4854–4870.

Duncan, R. O., Albright, T. D., & Stoner, G. R. (2000). Occlusion and the interpretation of visual motion: Perceptual and neuronal effects of context. *Journal of Neuroscience, 20*(15), 5885–5897.

Gattass, R., & Gross, C. G. (1981). Visual topography of striate projection zone (MT) in posterior superior temporal sulcus of the Macaque. *Journal of Neurophysiology, 46*(3), 621–638.

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

Maunsell, J. H. R., & Van Essen, D. C. (1983). Functional properties of neurons in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed and orientation. *Journal of Neurophysiology, 49*(5), 1127–1147.

Motter, B. C. (1994). Neural correlates of attentive selection for color or luminance in extrastriate area V4. *Journal of Neuroscience, 14*(4), 2178–89.

Movshon, J. A., Adelson, E. A., Gizzi, M., & Newsome, W. T. (1985). The analysis of moving visual patterns. In C. Chagas, R. Gattass, & C. G. Gross (Eds.), *Study group on pattern recognition mechanisms* (pp. 117–151). Vatican City: Pontifica Academia Scientiarum.

Newsome, W. T., Britten, K. H., & Movshon, J. A. (1989). Neuronal correlates of a perceptual decision. *Nature, 341*(6237), 52–54.

Newsome, W. T., & Pare, E. B. (1988). A selective impairment of motion perception following lesions of the middle temporal visual area (MT). *Journal of Neuroscience, 8*(6), 2201–2211.

Reynolds, J. H., Chelazzi, L., & Desimone, R. (1999). Competitive mechanisms subserve attention in macaque areas V2 and V4. *Journal of Neuroscience, 19*(5), 1736–1753.

Rodman, H. R., & Albright, T. D. (1987). Coding of visual stimulus velocity in area MT of the macaque. *Vision Research, 27,* 2035–2048.

Rodman, H. R., & Albright, T. D. (1989). Single-unit analysis of pattern-motion selective properties in the middle temporal visual area (MT). *Experimental Brain Research, 75,* 53–64.

Salzman, C. D., Britten, K. H., & Newsome, W. T. (1990). Cortical microstimulation influences perceptual judgements of motion direction. *Nature, 346*(6280), 174–177.

Shimojo, S., Silverman, G. H., & Nakayama, K. (1989). Occlusion and the solution to the aperture problem for motion. *Vision Research, 29*(5), 619–626.

Stoner, G. R., & Albright, T. D. (1992). Neural correlates of perceptual motion coherence. *Nature, 358*(6385), 412–414.

Stoner, G. R., & Albright, T. D. (1993). Image segmentation cues in motion processing: Implications for modularity in vision. *Journal of Cognitive Neuroscience, 5,* 129–149.

Swets, J. A., Green, D. M., Getty, D. J., & Swets, J. B. (1978). Signal detection and identification at successive stages of observation. *Perception and Psychophysics, 23*(4), 275–89.

Ungerleider, L. G., & Mishkin, M. (1979). The striate projection zone in the superior temporal sulcus of macaca mulatta: Location and topographic organization. *The Journal of Comparative Neurology, 188,* 347–366.

Van Essen, D. C., Maunsell, J. H. R., & Bixby, J. L. (1981). The middle temporal visual area in the macaque: Myeloarchitecture, connections, functional properties and topographic connections. *The Journal of Comparative Neurology, 199,* 293–326.

Wallach, H. (1935). Über visuell wahrgenommenr Bewegungsrichtung. *Psychologische Forschung, 20.* [About Visually Perceived Direction of Motion, *Phychological Research, 20,* 325–380.

Zeki, S. M. (1974). Cells responding to changing image size and disparity in the cortex of the rhesus monkey. *The Journal of Physiology, 242,* 827–841.

PERCEPTUAL ORGANIZATION IN VISION
BEHAVIORAL AND NEURAL PERSPECTIVES

Edited by

Ruth Kimchi
University of Haifa, Israel

Marlene Behrmann
Carnegie Mellon University

Carl R. Olson
Carnegie Mellon University
University of Pittsburgh

LEA
2003 LAWRENCE ERLBAUM ASSOCIATES, PUBLISHERS
Mahwah, New Jersey London

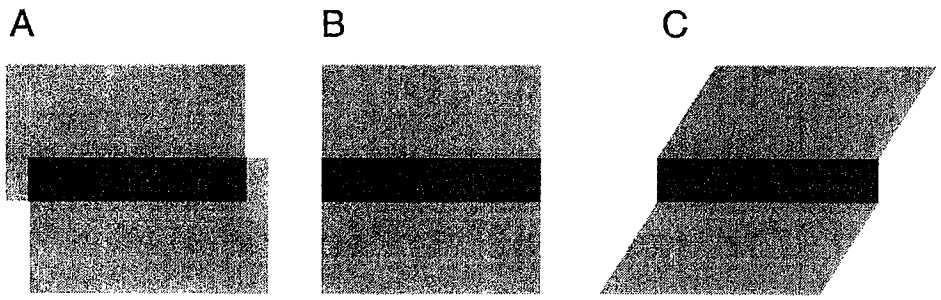

FIG. 9.1. Illustration of the influence of visual context on perception. Each of the three images displayed here contains a horizontal dark gray rectangle. Although these rectangles are physically identical, the surrounding features (the contexts) differ in the three images. As a result, the rectangle is perceptually attributed to different environmental causes in the three instances. (A) The rectangle appears to result from the overlap of two surfaces, one of which is transparent (e.g., a piece of tinted glass). (B) The rectangle appears to result from a variation in surface reflectance (e.g., a stripe painted across a large flat canvas). (C) The rectangle appears to result from partial shading of a surface (i.e., variation in the angle of the surface with respect to the source of illumination). These markedly different perceptual interpretations argue for the existence of different neuronal representations of the rectangle. These representations can only be identified in neurophysiological experiments if appropriate contextual cues are used for visual stimulation.

The different perceptual interpretations of the dark gray bars in Fig. 9.1, Panels A–C, follow from the different spatial contexts in which the bars appear. Such contextual influences, which are ubiquitous in normal visual experience, raise two important questions for neurobiologists: (1) At what stage of visual processing is contextual information incorporated to achieve neuronal representations of things perceived (scene-based representations), rather than local features of the retinal image (image-based representations)? (2) What neuronal mechanisms underlie the transformation from image-based to scene-based representations?

Our laboratory spent much of the past decade addressing these questions, primarily in the domain of visual motion processing (for review see Albright, 1993; Croner & Albright, 1999b; Stoner & Albright, 1993). In this review, we summarize two recent experiments from our laboratory. These experiments address the phenomenology and neuronal bases of two key elements of perceptual organization: feature interpretation and feature grouping.

FEATURE INTERPRETATION

When viewing any natural scene, the visual system interprets each feature—such as a corner, edge, patch, and so on—as belonging to a particular object. Fig. 9.2 illustrates the relevance of feature interpretation to motion processing. The scene in

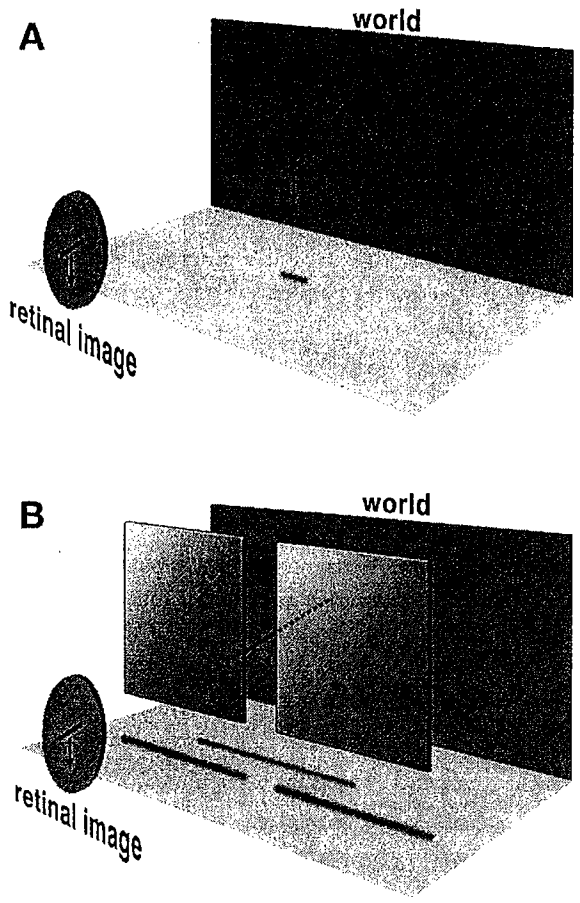

FIG. 9.2. Illustration of the contextual dependence of motion perception. Two simple visual scenes are shown, along with the retinal image motions that they give rise to. (A) This visual scene contains a single moving object (an elongated bar), which is oriented obliquely and moving directly down. The retinal image motion directly reflects the object motion. (B) This visual scene contains a single obliquely oriented bar, but in this case the bar is moving to the lower right (in a direction orthogonal to its orientation). The scene also contains two static opaque occluders, which block the observer's view of the two ends of the moving bar. The resulting retinal image motion in Panel B is identical to that caused by the scene in Panel A, although the visual scene motions are clearly different. The different contexts present in the two scenes—no surrounding features in Panel A, occluding panels present in Panel B—allow the different visual scenes to be perceived veridically.

Fig. 9.2A contains a single moving object—a sticklike figure—oriented obliquely and moving directly downward. The left-hand side of the figure illustrates the retinal image rendered by this simple scene. The scene in Fig. 9.2B also contains a single moving object—once again, a sticklike figure—obliquely oriented but moving toward the lower right corner of the scene. Unlike the scene in Fig. 9.2A, this scene also contains two opaque surfaces, each occluding one end of the moving