Chapter 3 Dynamics of Pattern Motion Computation

Matthew A. Smith, Najib Majaj, and J. Anthony Movshon

Abstract Early in visual processing neurons with small receptive fields can only signal the component of motion perpendicular to the orientation of the contour that passes through them (the "aperture problem"). A moving visual pattern with differently oriented contours can thus elicit neuronal responses that convey conflicting motion cues. To recover the true direction of motion of such a pattern, later visual areas must integrate the different motion cues over space and time. There is extensive evidence which suggests that this integration is not instantaneous – instead it occurs over time and causes profound changes in the perception of direction of motion of some complex moving patterns. To account for such temporal dynamics, previous studies have focused on a two-pathway model of motion perception: a fast pathway to account for the early percept, and a slow one to account for the late percept.

Neurons in macaque area MT are selective for the direction of motion of an object, and their responses appear to be connected directly to the perception of complex motion stimuli in the natural environment. In this chapter, we will discuss neurophysiological data from MT neurons which illustrate how the process of motion perception occurs dynamically. The responses of individual neurons in MT appear to reflect the process by which the primate visual system produces an initial estimate of motion direction and then refines it over time. We will argue that MT neuronal responses are consistent with a single pathway model of motion perception in which temporal dynamics emerge due to two factors: the contrast of elements in the pattern and the time required for the pattern computation.

M.A. Smith (\boxtimes)

Center for Neural Basis of Cognition, University of Pittsburgh, 4400 Fifth Avenue, Mellon Institute Room 115, Pittsburgh, PA, 15213, USA e-mail: masmith@cnbc.cmu.edu

3.1 Introduction

In the natural world, our visual experience is not static. Our eyes scan rapidly over the scene before us, fixing on a position for a few hundred milliseconds at a time before moving to the next location. Even within that short fixation, the image on our retina is rarely still – our own bodies and objects in the world around us are in nearly constant motion. Our perception of motion is similarly dynamic, changing over time on a scale of tens to hundreds of milliseconds. Nonetheless, most studies of visual cortex have measured the mean activity of neurons over a period of seconds. In recent years, however, the temporal dynamics of neural response in visual cortex have become subject to increased scrutiny. Perhaps not surprisingly, a number of recent studies have shown extensive changes in the response of visual cortical neurons over time. These dynamics appear to reflect the time course of multiple excitatory and suppressive influences which combine to produce a neuron's response. Comparison of the speed with which our psychophysical performance and physiological response unfold has proved to be an effective tool in understanding motion perception.

3.1.1 Temporal Dynamics in Primary Visual Cortex

Investigation of even the earliest stage of cortical visual processing has revealed significant dynamics in neuronal response. This includes modulation by stimuli which are confined to the receptive field and by those which extend well outside. A consideration of these findings in primary visual cortex (V1) is helpful in understanding the mechanisms of motion perception and the nature of perceptual effects. V1 neurons are tuned to the orientation of a stimulus within their receptive field (RF), and exhibit substantial time-dependent changes in that tuning (Ringach et al. 1997; Ringach et al. 2003; Smith et al. 2006). For small stimuli confined to the receptive field these effects tend to be extremely fast, often occurring with latency equal to or less than the excitatory response onset (Smith et al. 2006). The speed of such phenomena makes it likely that they are generated either by modification of the feedforward input to V1 neurons or very fast computations within the local circuitry in a cortical column.

Stimuli which extend beyond the receptive field, into the non-classical surround, also exert considerable influence on the responses of V1 neurons. In V1, neurons are modulated by a number of oriented stimuli which extend outside the receptive field. Various studies have shown that this effect arrives with some delay after the onset response for a number of stimuli, including fields of bars (Knierim and Van Essen 1992), oriented texture (Lamme 1995; Zipser et al. 1996; Lee et al. 1998), and sinusoidal gratings (Bair et al. 2003; Smith et al. 2006). The timing of surround suppression is one factor that has led most authors to conclude that it originates via feedback to V1 from extrastriate cortex. Angelucci et al. (2002), drawing on data from physiology and neuroanatomy, argued that the spatial scale of surround suppression is well matched to that of feedback circuits from extrastriate cortex to V1.

In searching for neural correlates of perceptual effects in V1, several studies have revealed the presence of dynamics in the response. These effects occur for a number of contextual stimuli presented to an awake animal performing a behavioral task, including curve-tracing (Roelfsema et al. 1998), figure-ground stimuli (Lamme 1995; Lee et al. 1998), illusory contours (Lee and Nguyen 2001), and shape-from-shading (Lee et al. 2002; Smith et al. 2007). The modulation of neuronal response in these paradigms occurs with greater delay than that found for extended iso-orientation stimuli. However, the pattern of neuronal response is similar: it remains normal for some time after the onset, and then after a delay the modulation due to the stimulus context becomes evident.

Finally, there is an additional factor which affects the dynamics of visual processing: the contrast of the stimulus. Specifically, low contrast targets are processed slower than high contrast ones (Albrecht 1995; Carandini et al. 1997; Gawne et al. 1996). This effect is distinct from observations of contextual modulation, in that the latency and magnitude of a neuron's response changes gradually with contrast, and this effect occurs for stimuli confined to the receptive field. Furthermore, the change in latency with contrast can be quite large, spanning up to 100 ms between the lowest and highest contrast stimuli.

3.1.2 Temporal Dynamics in Motion Perception

The relatively small receptive fields and proportion of direction-selective neurons in area V1 make it poorly suited for encoding complex moving stimuli. However, area MT of macaque visual cortex contains a high proportion of neurons which are selective for the direction of stimulus motion (Albright 1984; Movshon et al. 1985; Van Essen et al. 1981; Zeki 1974). Neurons in area MT have also been shown to play an important role in visual motion perception (Britten et al. 1992; Newsome and Paré 1988; Salzman and Newsome 1994). A significant portion of MT neurons encode the true velocity of a stimulus (Perrone and Thiele 2001; Priebe et al. 2003), whereas V1 neurons have independent spatial and temporary frequency responses (Tolhurst and Movshon 1975; Holub and Morton-Gibson 1981; Friend and Baker 1993). Similarly, many MT neurons are capable of decoding the true direction of motion of complex visual patterns such as a plaid stimulus (Fig. 3.2a), composed of two sinusoidal gratings with different orientations (Movshon et al. 1985; Rodman and Albright 1989), a behavior that is not present in V1 neurons that project to MT (Movshon and Newsome 1996). The responses of MT neurons to plaid patterns also vary in a manner consistent with the perceptual phenomenon of motion coherence (Stoner and Albright 1992). However, even though MT neurons have rather precise temporal response properties (Bair et al. 1994), the true direction of pattern motion is not represented in their initial responses. Instead, the encoding of pattern motion direction lags behind the initial estimate of direction by 50–75 ms (Pack and Born 2001; Smith et al. 2005). Taken together, these findings indicate that the detection of complex motion in the visual world is a property which emerges through computation in circuits within area MT.

The dynamics of neural response in area MT are paralleled by psychophysical results in humans. A number of studies have reported that manipulating the components of a complex moving stimulus affects the perception of human observers. When the speed of the component gratings of a plaid stimulus is unequal, observers perceive mostly component motion in brief presentations and pattern motion only after a delay (Yo and Wilson 1992). Similarly, short-latency ocular following responses in humans initially track the motion of a component grating, but later reflect the motion of the pattern (Masson and Castet 2002). A related effect occurs when the aspect ratio of an aperture around a grating stimulus is elongated observers first track the grating motion and over time are biased toward the aperture's long axis (Masson et al. 2000). Patterns composed of line segments have a similar effect on motion perception. With an array of line segments moving at various angles relative to their orientation, a pair of studies (Castet et al. 1993; Lorenceau et al. 1993) reported that observers were biased by the orientation of the line segments in short observation windows (100–200 ms), but over time they recognize the true motion direction. The initiation of smooth-pursuit eye movements in humans shows a similar bias for stimuli composed of line segments (Masson and Stone 2002), diamonds (Wallace et al. 2005), and a combination of first and second-order motion cues (Lindner and Ilg 2000). The common finding in all of these studies is that observers tend to show an initial bias toward the orientation of components of a pattern, but over the span of 100 ms or more, tend to perceive the true motion direction. The dynamics evident in these results have led to a number of models of how the visual system computes pattern motion. We will consider each of these models in turn.

3.2 Models of Pattern Motion Detection

Consideration of the data from physiological and psychophysical studies has revealed that motion information is processed in at least two stages. The first, likely located in primary visual cortex (V1), extracts basic information (such as orientation) about simple moving patterns from a local region of space. The second stage computes information about the true direction and speed of complex moving patterns by combining inputs from the first stage. Models which strive to explain motion perception typically reflect this two-stage processing in their instantiation.

One such model linearly combines the signals from nonlinear V1 subunits (Heeger et al. 1996; Simoncelli and Heeger 1998), a so-called linear–nonlinear ("L–N") model. This simple model is able to capture many of the properties of direction-selective neurons in macaque area MT, but cannot adequately account for pattern direction selectivity (Simoncelli and Heeger 1998; Mante 2000). A modification of this basic structure, a cascaded L–N model in which the second stage acts on signals from a population of direction-selective units, can accurately decode the motion of complex patterns while maintaining fidelity to the known cortical architecture (Rust et al. 2006). In the cascade model (Fig. 3.1), a stimulus passes through



MT functional model

Fig. 3.1 The cascade model

a population of direction-selective V1 neurons and is divisively normalized. The outputs of these model cells feed into a MT neuron which computes their linear weighted sum. The result is converted into firing rate by a nonlinear function, simulating the effect of spike threshold and any additional nonlinear effects which occur post-summation. In the framework of this model, pattern selectivity arises from the recurrent circuit which combines V1 inputs to produce MT neuronal responses. If such a network takes time to stabilize, the selectivity of individual neurons would change over time – first reflecting the simple direction selectivity of input neurons, and later evolving pattern selective responses. One aspect of the cascade model, an orientation-tuned normalization mechanism, may reflect suppressive input from outside the classical receptive field of V1 neurons. This surround suppression is known to occur with some delay after response onset (Bair et al. 2003; Smith et al. 2006), and may lead to the delay in pattern motion computation observed in MT neurons (Smith et al. 2005).

Another class of model, which separately analyzes the contour and terminator information present in a scene (Shimojo et al. 1989; Grossberg and Mingolla 1993; Lorenceau et al. 1993), has also been proposed as a means of decoding motion in complex patterns. A version of this approach uses two parallel pathways (Fourier and non-Fourier), the outputs of which are combined to compute pattern motion (Wilson et al. 1992; Wilson and Kim 1994; Löffler and Orbach 1999). However, the cortical pathways which underlie this model are unknown, and studies of the proposed candidate areas (V2 and V3) do not suggest an important contribution to pattern selectivity (Gegenfurtner et al. 1997; Levitt et al. 1994).

A third approach proposes that neural networks separately process the ambiguous and unambiguous portions of the scene, with the unambiguous locations "filling in" over time (Hildreth 1984; Beutter and Stone 1998). Neural models using recurrent (Lidén and Pack 1999) or feedback (Chey et al. 1997) circuits to decode pattern motion have implemented this proposal. If such a model is tuned so that early responses reflect feedforward signals, and later responses are shaped by

recurrent or feedback connections, then the dynamics of pattern motion perception can be replicated.

These three classes of model may provide us with some insight into the dynamics of pattern motion computation. By adjusting the latency or dynamics of the two stages or pathways in each model, it is possible to generate a pattern motion detection system with dynamics which are similar to those shown in experimental studies. It is clear that further exploration of neurophysiological responses is necessary to distinguish the neural mechanisms clearly. We will now describe one series of experiments which aims to explore these mechanisms, and explain the neural basis for our changing perception of complex moving patterns over time.

3.3 Responses of MT Neurons to Plaids

Area MT in the extrastriate cortex of the macaque contains a high proportion of directionally selective neurons (Albright 1984; Movshon et al. 1985; Van Essen et al. 1981; Zeki 1974) and plays an important role in the perception of moving patterns (Britten et al. 1992; Newsome and Paré 1988). When presented with a drifting sinusoidal grating stimulus, the vast majority of MT neurons respond in a direction selective manner (Fig. 3.2b, left). Plaid stimuli, obtained by adding two sinusoidal gratings with different orientations (Fig. 3.2a, right), have been used to demonstrate an important property of some MT neurons which is not present at earlier stages of motion processing. When presented with a plaid stimulus, a directionally selective neuron might respond only to the direction of the component gratings (solid line in Fig. 3.2b, right plot), or it might respond to the true direction of motion of the plaid stimulus (dashed line in Fig. 3.2b, right plot). The former behavior is termed component direction selectivity (CDS) and the latter pattern direction selectivity (PDS).



Fig. 3.2 Pattern and component selectivity (modified from Smith et al. 2005)

The classification of a neuron as CDS or PDS is made by comparing its actual tuning curve to a plaid, with two predictions (Fig. 3.2b, right plot) made based on its direction tuning to a single grating. In a population of MT neurons (Fig. 3.2c), 25% were classified as PDS (*white circles*) and 41% were classified as CDS (*gray circles*), with the remainder unclassed (Smith et al. 2005). The solid lines indicate significance boundaries for the classification of neurons as PDS or CDS. V1 neurons signal only the direction of motion of the component gratings (CDS) and not the true pattern direction (Movshon et al. 1985; Movshon and Newsome 1996). This is also true of V1 neurons which project directly to MT (Movshon and Newsome 1996), which is consistent with the idea that pattern motion is computed by circuits within MT.

3.4 Dynamics of MT Neuronal Response

There is evidence from both physiology and psychophysics that the neural representation of complex patterns evolves over tens to hundreds of milliseconds (Pack and Born 2001; Kooi et al. 1992; Lorenceau et al. 1993; Yo and Wilson 1992; Masson and Castet 2002). Since MT neurons are known to play a role in the perception of complex moving patterns, and human observers of plaid stimuli appear to refine their estimate of the direction over time, MT is a natural location in which to look for dynamics of response to such patterns.

Figure 3.3a-d contain scatter plots of pattern and component correlation, computed in the same way as the one shown in Fig. 3.2c. Each panel shows data taken from a small window of time cut out from the full stimulus period. Each point represents one neuron, and the points are colored to indicate the selectivity of that neuron over the entire stimulus period (CDS neurons are gray circles, PDS neurons are *white circles*, and unclassed neurons are *black circles*). In the top panel (Fig. 3.3a), only the period from 30 to 50 ms after stimulus onset is included. In this response window, before the onset latency for many MT neurons, there is little or no significant tuning for this measure. In Fig. 3.3, the response window in each row includes an additional 20 ms of data. CDS selective behavior is evident (as indicated by the gray circles that have already crossed the significance line) only 70 ms after stimulus onset (Fig. 3.3b). PDS neurons, however, take longer to show their characteristic behavior. Some neurons reach significance by 90 ms after stimulus onset (Fig. 3.3c), while many others take longer - up to 110 ms (Fig. 3.3d). At this time, 110 ms after stimulus onset, most of the CDS neurons but less than half of the PDS neurons have responded in a way consistent with their final tuning. In the right column (Fig. 3.3e-h), the same analysis is shown using a sliding window. Comparison of the rows reveals that this phenomenon is not due to the reduced noise that comes from averaging over a longer response window. Instead, these scatter plots demonstrate that PDS neurons lag behind CDS neurons in the time it takes them to show their characteristic response.



Fig. 3.3 Evolution of pattern and component selectivity for individual neurons

3 Dynamics of Pattern Motion Computation

Figure 3.4a shows the evolution of PDS and CDS behavior for the population averages of the three classes of neurons from Fig. 3.2c (Smith et al. 2005). For cells which are eventually labeled as PDS, CDS, or unclassed (based on their response over the full stimulation period), the three lines show the evolution of their pattern and component correlation values over time, starting with the stimulus onset (time after stimulus onset is indicated with the numbers and connected lines). The CDS neurons (*dashed gray* line) cross the significance threshold much earlier in the



Fig. 3.4 Evolution of pattern and component selectivity in the population average response (modified from Smith et al. 2005)

response than the PDS neurons (*solid gray* line). The difference between the pattern and component indices is shown in Fig. 3.4b, to produce an index of "paternness" for PDS cells (*solid gray* line) and "componentness" for CDS cells (*dashed gray* line). When examining the times at which these two populations cross the significance threshold (horizontal black line), there is a difference of approximately 60–70 ms between the average CDS cell and the average PDS cell.

CDS neurons therefore develop their characteristic response tuning much earlier (60–65 ms) than PDS neurons (125–130 ms), a trend which is evident in the population and also in the responses of individual neurons (Smith et al. 2005). This additional time for pattern direction selectivity to become manifest is considerable, and suggests that circuits more complex than a simple feed-forward network are involved in the computation of pattern motion.

3.5 Relationship Between Bar and Plaid Stimuli

Bar textures moving obliquely to their orientation change apparent direction at low speeds and contrasts. Initially they seem to be moving perpendicular to the orientation of the bar, and over the course of 200–300 ms a human observer's perception tends to shift to the true direction of motion (Castet et al. 1993; Lorenceau et al. 1993). Neurons in macaque area MT exhibit behavior which is analogous – their initial response is based on motion perpendicular to the bar orientation, and their preference later shifts toward the direction in which the terminators move (Pack and Born 2001). These results are consistent with models in which the contours and terminators in a scene are analyzed separately, on different timescales – a rapid signal related to contours or edges and a slower signal related to terminators (Shimojo et al. 1989; Grossberg and Mingolla 1993; Lorenceau et al. 1993; Wilson et al. 1992; Wilson and Kim 1994; Löffler and Orbach 1999).

These models would be able to produce results similar to physiological data through a transition between the two motion signals. If the terminator-related motion signal is processed separately and more slowly than signals related to contours, it is possible to explain physiological and psychophysical observations of dynamic changes in motion perception. There is an alternative interpretation, however, which does not rely on a separate pathway for processing terminator motion. Instead, it may be that contours and terminators only appear to be processed separately due to the well known effect of contrast on visual processing: lower contrast targets are processed slower than high contrast ones (Albrecht 1995; Carandini et al. 1997; Gawne et al. 1996). In order to assess this proposal, we first have to consider a frequency analysis of the bar texture stimuli.

Figure 3.5a shows standard line, dot and bar stimuli in the top row. The motion of the first stimulus (*left*), a series of lines behind an aperture (a rectangular grating), is ambiguous. Motion of this stimulus in any direction is interpreted by an observer (and by MT neurons) as being perpendicular to the orientation of the lines. When there is motion of the second stimulus (*middle*) - a field of dots - it is unambiguous. The third stimulus, a bar texture (*right*), contains elements common



Fig. 3.5 Grating and bar stimuli

to both lines and dots. It is this stimulus which elicits a dynamically changing perception in observers (Castet et al. 1993; Lorenceau et al. 1993) and response in MT neurons (Pack and Born 2001).

The Fourier Transform decomposes any image into a sum of sinusoids of various amplitudes, spatial frequencies, and orientations. In the case of lines, dots, and bars, the Fourier Transform can provide a concrete basis for understanding the commonalities and differences between the stimuli. The first stimulus (*left*), a rectangular grating, has a Fourier amplitude spectrum composed of a sum of sinusoidal gratings of the same orientation but a range of spatial frequencies. In contrast, the dots (middle) have a broad amplitude spectrum, consisting of a sum of sinusoidal gratings across a wide range of orientations and spatial frequencies. For both the rectangular grating and dots, however, all the components have similar amplitude (contrast). Although the bar texture (*right*) may at first appear to consist of contours of only one orientation, examination of its Fourier amplitude spectrum reveals that this is not the case. Although it shares features of the lines (a dominant axis of power at one orientation) and dots (broad power), the bar texture's spectrum differs in an important way - the amplitude (contrast) of the components depends on their orientation. Components parallel to the orientation of the bar texture have the highest amplitude, while those at oblique orientations have lower contrast. Changing the contrast of these oblique components has an interesting effect -a decrease in amplitude lengthens the bars until they connect and form continuous lines, while increasing shortens the bar length until they approximate dots.

The bottom row in Fig. 3.5b contains filtered analogs of the corresponding stimuli in the top row. These filtered textures contain only the fundamental spatial frequency components of the original images in Fig. 3.5a. The line stimulus can be approximated by a single grating (*left*) and the dot stimulus by a type of plaid composed of four gratings of equal contrast (*middle*). The filtered bar texture

(*right*) is produced by combining the same four constituent gratings but with unequal contrast: a high contrast grating parallel to the bar texture orientation, and three low contrast gratings at $+45^{\circ}$, $+90^{\circ}$ and $+135^{\circ}$ relative to the high contrast grating. The filtered bar image has no obvious terminators, but nonetheless retains the essential structure of the original bar image. Thus, it is a good stimulus with which to test our hypothesis about the role of contrast in processing bar textures.

3.6 Response to Filtered Bar Textures

If the filtered bar texture shown on the bottom right of Fig. 3.5 is indeed a good approximation to the corresponding bar texture directly above it, then MT neurons should have dynamically evolving responses to this stimulus. Figure 3.6a shows the temporal evolution of direction selectivity in a population of MT neurons to



Fig. 3.6 Dynamics of direction selectivity to filtered textures

the three filtered stimuli described above (Majaj 2006). In order to compute the population average, the responses are adjusted so that both the direction preference and onset latency are aligned. When all four components of the filtered stimuli are of equal contrast, the resulting filtered dot texture appears to move to the right, and the neuronal population exhibits consistent direction tuning (*thick black line*). With a single component (a grating), the population of neurons responds prefers the direction orthogonal to the orientation (*thin black line*). These two results are expected based on previous physiological studies of MT and psychophysical studies of human motion perception. For the filtered bar texture, possessing components of different contrast, the direction preference does not remain constant over time. Instead, there is a distinct change in the direction preference over a period extending to 120 ms after stimulus onset (*medium black line*). The early response resembles the direction preference to a single grating; it changes over the next 50 ms, and by 120 ms it stabilizes to resemble the direction preference dots.

The change in direction selectivity for filtered bar textures resembles the psychophysical findings of Lorenceau et al. (1993) and physiological results of Pack and Born (2001). Figure 3.6b shows the evolution of direction selectivity for filtered bar textures (middle line from Fig. 3.6a) compared with that for unfiltered bar textures (data replotted from Pack and Born 2001). The dynamics observed in response to the filtered bar texture are remarkably similar to those obtained in response to the texture itself, even though the filtered texture is composed of four sinusoidal gratings and contains no obvious terminators. This suggests that a common mechanism might underlie the dynamics in response to these two stimuli.

3.7 Effects of Contrast on Response Dynamics

The similarity between these two results might be explained by a well known phenomenon in vision. The contrast of a visual stimulus has a powerful effect on the speed of visual processing: neurons respond to low contrast stimuli with a long latency, but high contrast stimuli are processed with a short latency (Albrecht 1995; Carandini et al. 1997; Gawne et al. 1996). The filtered bar texture stimulus is composed of gratings with different contrast – a high contrast grating parallel to the bar texture orientation, and three low contrast oblique gratings. If the different components of the filtered bar texture are processed with different latencies, then the response to the combined stimulus might be expected to exhibit dynamics.

Stimulus contrast affects response magnitude and latency in MT neurons. Figure 3.7 shows this effect for the responses of a single MT neuron. Each row of the figure shows a raster plot of responses to repeated presentations of a drifting grating stimulus in the preferred direction, ranging from 100% contrast (top row) to 10% contrast (bottom row). A histogram, binned at 1 ms precision, of the firing



Fig. 3.7 Effect of contrast on response latency

rate over time is shown with a *gray line* in each row. The response latency of this neuron, measured from the response onset or the time to reach peak firing rate, grows as contrast decreases (from the top to bottom rows). In this example neuron, the onset latency changes from approximately 65 to 110 ms as the contrast decreases. The trend evident in the responses shown for the single neuron in Fig. 3.7 is also observed in a larger population of MT neurons (Thiele and Hoffman 1996; Thiele et al. 1999; Majaj 2006). A delay of 100 ms is typical between onset latency at low and high contrast.

In the filtered bar texture, the component gratings may be processed by the visual system with different latencies due to their different contrasts. The highest contrast component, parallel to the orientation of the comparable bar texture, would be processed first. The other components - lower in contrast and at oblique angles relative to the primary component - would be processed with some delay. As an MT neuron integrates the information in all of these components, its direction preference will change over time. The early responses will be dominated by the high contrast grating, but as the lower contrast oblique gratings are processed the later responses will reflect this information.

3.8 Conclusions

In natural vision, our experience of the world is rich with temporal dynamics. Our visual system has evolved to be able to evaluate and interpret this information with the speed necessary to make fast judgments based on our perceptual experience. Nonetheless, the integration necessary to determine the motion direction of a complex visual pattern is not instantaneous. In this chapter, we have described physiological evidence that neurons in macaque area MT have response dynamics which evolve over the first 100–200 ms after a visual stimulus appears. The temporal profile and directional selectivity of these responses parallels results from a well established psychophysical literature based on experiments in human observers.

A number of models have been proposed to explain the dynamics observed in physiological and psychophysical studies. The most common approach has been to propose separate pathways for analyzing different visual features (terminators and contours) – essentially, parallel one- and two-dimensional motion analysis. Here, we have reasoned from data and models that one-dimensional orientation-selective mechanisms can account for these experimental observations. This is not to imply that two-dimensional features are unimportant for visual processing, but rather that a separate pathway for analysis of such features is not necessary to explain dynamics in motion perception. A linear–nonlinear cascade, which incorporates a number of physiologically realistic processes into a functional model, is an alternative approach which provides a good fit to the neuronal and perceptual data.

3.9 Supplementary materials (CD-ROM)

Movie 1 Field of moving upright bars (file *«*3_*M1_Bars0.avi »*) A field of bars moving perpendicular to their orientation (all of the motion cues in the stimulus are consistent). Videos 1 & 2 are modeled after the stimuli used by Pack and Born (2001).

Movie 2 Field of moving tilted bars (file $\ll 3_M2_Bars45.avi \gg$) Moving bars with an orientation of $+45^\circ$ relative to the direction of motion. This is the moving version of the stimulus shown in the right panel of Fig. 3.5a. In this case, the local motion signal along the edges of the bars is in conflict with the true direction of motion.

Movie 3 Grating motion (see Fig. 3.5) (file $(3_M3_Filteredlines.avi)$) A single drifting grating stimulus at +45° (Fig. 3.5b, left panel). Because there is only a single orientation present in the stimulus, and no access to terminators (the "aperture problem"), the direction of motion is perceived to be perpendicular to the orientation of the grating.

Movie 4 Filtered dot texture (see Fig. 3.5) (file $(3_M4_Filtereddots.avi)$) The filtered analog of a dot texture, this is a type of plaid stimulus (Fig. 3.5b, middle panel) composed of four gratings of equal contrast. The grating orientations are $+0^\circ$,

 $+45^{\circ}$, $+90^{\circ}$ and $+135^{\circ}$ relative to the bar texture shown in Movie 2. The result is a filtered dot texture similar to the dot image shown in Fig. 3.5a (middle panel).

Movie 5 Filtered bar texture (see Fig. 3.5) (file $<3_M5_Filteredbars.avi >>$) This filtered bar texture (Fig. 3.5b, right panel) is composed of the same four gratings as in Movie 4, but with unequal contrast: a high contrast grating parallel to the bar texture orientation, and three low contrast gratings at +45°, +90° and +135° relative to the high contrast grating.

Movie 6 Filtered bar texture and grating components (file «*3_M6_Texture-components.avi* »). This video shows the filtered bar texture (from Video 5) decomposed into its four constituent gratings and then re-assembled.

References

- Albrecht DG (1995) Visual cortex neurons in monkey and cat: effect of contrast on the spatial and temporal phase transfer functions. Vis Neurosci 12:1191–1210
- Albright TD (1984) Direction and orientation selectivity of neurons in visual area MT of the macaque. J Neurophysiol 52:1106–1130
- Angelucci A, Levitt JB, Walton EJ, Hupé JM, Bullier J, Lund JS (2002) Circuits for local and global signal integration in primary visual cortex. J Neurosci 22:8633–8646
- Bair W, Cavanaugh JR, Movshon JA (2003) Time course and time-distance relationships for surround suppression in macaque V1 neurons. J Neurosci 23:7690–7701
- Bair W, Koch C, Newsome W, Britten K (1994) Power spectrum analysis of bursting cells in area MT in the behaving monkey. J Neurosci 14:2870–2892
- Beutter BR, Stone LS (1998) Human motion perception and smooth eye movements show similar directional biases for elongated apertures. Vision Res 38:1273–1286
- Britten KH, Shadlen MN, Newsome WT, Movshon JA (1992) The analysis of visual motion: a comparison of neuronal and psychophysical performance. J Neurosci 12:4745–4765
- Carandini M, Heeger DJ, Movshon JA (1997) Linearity and normalization in simple cells of the macaque primary visual cortex. J Neurosci 17:8621–8644
- Castet E, Lorenceau J, Shiffrar M, Bonnet C (1993) Perceived speed of moving lines depends on orientation, length, speed and luminance. Vision Res 33:1921–1936
- Chey J, Grossberg S, Mingolla E (1997) Neural dynamics of motion grouping: From aperture ambiguity to object speed and direction. J Opt Soc Am A 14:2570–2594
- Friend SM, Baker CL Jr (1993) Spatio-temporal frequency separability in area 18 neurons of the cat. Vision Res 33:1765–1771
- Gawne TJ, Kjaer TW, Richmond BJ (1996) Latency: another potential code for feature binding in striate cortex. J Neurophysiol 76:1356–1360
- Gegenfurtner KR, Kiper DC, Levitt JB (1997) Functional properties of neurons in macaque area V3. J Neurophysiol 77:1906–1923
- Grossberg S, Mingolla E (1993) Neural dynamics of motion perception: direction fields, apertures, and resonant grouping. Percept Psychophys 53:243–278
- Heeger DJ, Simoncelli EP, Movshon JA (1996) Computational models of cortical visual processing. Proc Natl Acad Sci USA 93:623–627
- Hildreth EC (1984) The measurement of visual motion. MIT Press, Cambridge, MA
- Holub RA, Morton-Gibson M (1981) Response of visual cortical neurons of the cat to moving sinusoidal gratings: response-contrast functions and spatiotemporal interactions. J Neurophysiol 46:1244–1259
- Knierim JJ, Van Essen DC (1992) Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. J Neurophysiol 67:961–980

- Kooi FL, DeValois KK, Switkes E, Grosof DH (1992) Higher-order factors influencing the perception of sliding and coherence of a plaid. Perception 21:583–598
- Lamme VAF (1995) The neurophysiology of figure-ground segregation in primary visual cortex. J Neurosci 15:1605–1615
- Lee TS, Mumford D, Romero R, Lamme VAF (1998) The role of primary visual cortex in higher level vision. Vision Res 38:2429–2454
- Lee TS, Nguyen M (2001) Dynamics of subjective contour formation in early visual cortex. Proc Natl Acad Sci USA 98:1907–1911
- Lee TS, Yang CF, Romero RD, Mumford D (2002) Neural activity in early visual cortex reflects behavioral experience and higher-order perceptual saliency. Nat Neurosci 5:589–597
- Levitt JB, Kiper DC, Movshon JA (1994) Receptive fields and functional architecture of macaque V2. J Neurophysiol 71:2517–2542
- Lidén LH, Pack CC (1999) The role of terminators and occlusion cues in motion integration and segmentation: a neural network model. Vision Res 39:3301–3320
- Lindner A, Ilg UJ (2000) Initiation of smooth-pursuit eye movements to first-order and secondorder motion stimuli. Exp Brain Res 133:450–456
- Löffler G, Orbach HS (1999) Computing feature motion without feature detectors: a model for terminator motion without end-stopped cells. Vision Res 39:859–871
- Lorenceau J, Shiffrar M, Wells N, Castet E (1993) Difference motion sensitive units are involved in recovering the direction of moving lines. Vision Res 33:1207–1217
- Majaj N (2006) Spatial and temporal integration of motion signals in area MT. Thesis, New York University
- Mante V (2000) Testing models of cortical area MT. Thesis, Institute of Neuroinformatics, ETH, University of Zurich
- Masson GS, Castet E (2002) Parallel motion processing for the initiation of short-latency ocular following in humans. J Neurosci 22:5149–5163
- Masson GS, Rybarczyk Y, Castet E, Mestre DR (2000) Temporal dynamics of motion integration for the initiation of tracking eye movements at ultra-short latencies. Vis Neurosci 17:753–767
- Masson GS, Stone LS (2002) From following edges to pursuing objects. J Neurophysiol 88:2869–2873
- Movshon JA, Adelson EH, Gizzi MS, Newsome WT (1985) The analysis of visual moving patterns. In: Chagas C, Gattass R, Gross C (eds) Pattern recognition mechanisms. Springer, New York, pp 117–151
- Movshon JA, Newsome WT (1996) Visual response properties of striate cortical neurons projecting to area MT in macaque monkeys. J Neurosci 16:7733–7741
- Newsome WT, Paré EB (1988) A selective impairment of motion perception following lesions of the middle temporal area MT. J Neurosci 8:2201–2211
- Pack CC, Born RT (2001) Two-dimensional substructure of MT receptive fields. Nature 409:1040–1042
- Perrone JA, Thiele A (2001) Speed skills: measuring the visual speed analyzing properties of primate MT neurons. Nat Neurosci 4:526–532
- Priebe NJ, Cassanello CR, Lisberger SG (2003) The neural representation of speed in macaque area MT/V5. J Neurosci 23:5650–5661
- Ringach DL, Hawken MJ, Shapley R (1997) The dynamics of orientation tuning in the macaque monkey striate cortex. Nature 387:281–284
- Ringach DL, Hawken MJ, Shapley R (2003) Dynamics of orientation tuning in macaque v1: the role of global and tuned suppression. J Neurophysiol 90:342–352
- Rodman HR, Albright TD (1989) Single-unit analysis of pattern-motion selective properties in the middle temporal visual area (MT). Exp Brain Res 75:53–64
- Roelfsema PR, Lamme VA, Spekreijse H (1998) Object-based attention in the primary visual cortex of the macaque monkey. Nature 395:376–381
- Rust N, Mante V, Simoncelli EP, Movshon JA (2006) How MT cells analyze the motion of visual patterns. Nat Neurosci 9:1421–1431

- Salzman CD, Newsome WT (1994) Neural mechanisms for forming a perceptual decision. Science 264:231–237
- Shimojo S, Silverman GH, Nakayama K (1989) Occlusion and the solution to the aperture problem for motion. Vision Res 29:619–626
- Simoncelli EP, Heeger DJ (1998) A model of neuronal responses in visual area MT. Vision Res 38:743–761
- Smith MA, Bair W, Movshon JA (2006) Dynamics of suppression in macaque primary visual cortex. J Neurosci 26:4826–4834
- Smith MA, Kelly RC, Lee TS (2007) Dynamics of response to perceptual pop-out stimuli in macaque V1. J Neurophysiol 98:3436–3449
- Smith MA, Majaj NJ, Movshon JA (2005) Dynamics of motion signaling by neurons in macaque area MT. Nat Neurosci 8:220–228
- Stoner GR, Albright TD (1992) Neural correlates of perceptual motion coherence. Nature 358:412-414
- Thiele A, Distler C, Hoffman KP (1999) Decision-related activity in the macaque dorsal visual pathway. Eur J Neurosci 11:2044–2058
- Thiele A, Hoffman KP (1996) Reaction time and time course of neuronal responses in MT and MST at different stimulus contrasts. Perception 25:ECVP Abstract Supplement
- Tolhurst D, Movshon J (1975) Spatial and temporal contrast sensitivity of striate cortical neurones. Nature 257:674-675
- Van Essen DC, Maunsell JHR, Bixby JL (1981) The middle temporal visual area in the macaque: myeloarchitecture, connections, functional properties and topographic organization. J Comp Neurol 199:293–326
- Wallace JM, Stone LS, Masson GS (2005) Object motion computation for the initiation of smooth pursuit eye movements in humans. J Neurophysiol 93:2279–2293
- Wilson HR, Ferrera VP, Yo C (1992) A psychophysically motivated model for two-dimensional motion perception. Vis Neurosci 9:79–97
- Wilson HR, Kim J (1994) Perceived motion in the vector sum direction. Vis Neurosci 34:1835–1842
- Yo C, Wilson HR (1992) Perceived direction of moving two-dimensional patterns depends on duration, contrast and eccentricity. Vision Res 32:135–147
- Zeki SM (1974) Functional organization of a visual area in the posterior bank of the superior temporal sulcus of the rhesus monkey. J Physiol (Lond) 236:549–573
- Zipser K, Lamme VA, Schiller PH (1996) Contextual modulation in primary visual cortex. J Neurosci 16:7376–7389