Cortical processing of second-order motion

ISABELLE MARESCHAL* AND CURTIS L. BAKER, JR.

McGill Vision Research Unit, Department of Ophthalmology, McGill University, 687 Pine Avenue West (H4 14),
Montreal, PQ, H3A 1A1, Canada

(Received July 6, 1998; Accepted November 19, 1998)

Abstract

Neurons in the mammalian visual cortex have been found to respond to second-order features which are not defined by changes in luminance over the retina (Albright, 1992; Zhou & Baker, 1993, 1994, 1996; Mareschal & Baker, 1998a,b). The detection of these stimuli is most often accounted for by a separate nonlinear processing stream, acting in parallel to the linear stream in the visual system. Here we examine the two-dimensional spatial properties of these nonlinear neurons in area 18 using envelope stimuli, which consist of a high spatial-frequency carrier whose contrast is modulated by a low spatial-frequency envelope. These stimuli would fail to elicit a response in a conventional linear neuron because they are designed to contain no spatial-frequency components overlapping the neuron’s luminance defined passband. We measured neurons’ responses to these stimuli as a function of both the relative spatial frequencies and relative orientations of the carrier and envelope. Neurons’ responses to envelope stimuli were narrowband to the carrier spatial frequency, with optimal values ranging from 8- to 30-fold higher than the envelope spatial frequencies. Neurons’ responses to the envelope stimuli were strongly dependent on the orientation of the envelope and less so on the orientation of the carrier. Although the selectivity to the carrier orientation was broader, neurons’ responses were clearly tuned, suggesting that the source of nonlinear input is cortical. There was no fixed relationship between the optimal carrier and envelope spatial frequencies or orientations, such that nonlinear neurons responding to these stimuli could perhaps respond to a variety of stimuli defined by changes in scale or orientation.

Keywords: Second-order, Cortical processing, Motion

Introduction

Despite a large body of evidence for nonlinearities arising in cortical processing (see Shapley & Lennie, 1985; Bonds 1992 for review), the selectivity of neurons in the primary visual cortex to stimulus attributes is conventionally explained in terms of the linear summation of luminance inputs falling over their receptive fields (Hubel & Weisel, 1977; Movshon et al., 1978; DeValois et al., 1982; Pollen & Ronner, 1982; Carandini et al., 1999). In this scheme, neurons are likened to linear filters which are selectively tuned to the spatiotemporal frequencies and orientations composing the stimulus.

Recently, a class of stimuli have been designed whose motion or orientation are invisible to linear spatiotemporal filters yet are easily detected by human observers (Chubb & Sperling, 1988; Cavanagh & Mather, 1989). These stimuli are referred to as second-order or non-Fourier because the spatiotemporal variations underlying their detection must be derived from a second-order statistic of the image. These stimuli have been postulated to mimic second- or higher order cues in the natural environment underlying visual

*Present address: Center for Neural Science, New York University, 4 Washington Place, New York, NY 10003.
Correspondence and reprint requests to: Curtis Baker, McGill Vision Research Unit, 687 Pine Avenue West (H4, 14), Montreal, PQ, H3A 1A1, Canada.

In the visual cortex, neurons have been identified which respond to second-order stimuli in a highly selective manner (Albright, 1992; Zhou & Baker, 1993, 1994, 1996; Mareschal & Baker, 1998a,b). Some of these experiments used second-order envelope stimuli which consist of a stationary high spatial-frequency grating (“carrier”) whose contrast is modulated by a low spatial-frequency grating (“envelope”). Since the neurons’ responses were contingent on the stimulus parameters, the authors concluded that an early nonlinearity preceding linear neuronal filtering could not account for the detection of these stimuli. Indeed, a mechanism which nonlinearly transformed the stimulus prior to frequency filtering could introduce energy at the spatiotemporal frequency (for example, retinal processing has been postulated to introduce “distortion products”; MacLeod et al., 1992). However, such an early nonlinearity would operate independently of the spatial frequency of the carrier. Previous studies found both a narrow spatial-frequency dependence on the carrier, as well as spatial and temporal differences between responses to the envelope and a luminance grating (Zhou & Baker, 1993, 1994, 1996; Mareschal & Baker, 1998a,b). Neither of these findings can be reconciled with an early nonlinearity hypothesis.

To account for the selective detection of second-order stimuli, a parallel nonlinear processing stream acting independently of a
conventional linear stream has been incorporated into most models of visual processing (Graham et al., 1992; Wilson et al., 1992; Zhou & Baker, 1993; Ledgeway & Smith, 1994; Nishida et al., 1997; but see Johnston et al., 1992). In this scheme, the nonlinear stream consists of two filtering stages with an interposed nonlinearity (referred to as “filter-rectify-filter” models). The first filtering stage provides selectivity to the carrier; its rectified output then feeds into a second filtering stage which mediates selectivity to attributes of the envelope, such as its direction of motion or orientation.

Psychophysical experiments examining the spatial characteristics of second-order detection using either static or dynamic stimuli have produced conflicting results. Using static stimuli Jamar and Koenderink (1985) measured detection thresholds for amplitude modulated noise and found no optimal spatial-frequency relationship between the carrier and envelope. Dakin and Mareschal (1998) found similar results when they measured detection thresholds for contrast modulated filtered noise as a function of the spatial frequency and orientation of the carrier. Although they found no optimal spatial relationship between carrier and envelope, their results suggested that the first filtering stage was band-pass in both spatial frequency and orientation. Similarly, Cropper (1998) found no carrier spatial-frequency dependence for the discrimination of amplitude modulated patterns. Using an adaptation paradigm, Langley et al. (1996) suggested preferential orientational relationships between the carrier and envelope. Their results also suggested a constant optimal envelope frequency which was largely independent of the carrier frequencies they tested. However, these results are inconsistent with Sutter et al. (1995) who examined spatial-frequency relationships between the carrier and envelope over a smaller range of frequencies and suggested a characteristic optimal ratio of carrier to envelope spatial frequencies (approximately 16-fold).

Using moving stimuli, Werkhoven et al. (1994) found that apparent motion could be elicited between two stimulus frames containing different spatial frequencies and orientations. They proposed the use of a broad “texture grabber” having spatial and orientational responses feeding into second-stage motion extractors. This result is in accord with Boulton and Baker (1994) who found that motion could be elicited in certain conditions between patterns differing in the spatial frequency or orientation content of the carrier. However, Cropper and Badcock (1995) found that the perceived direction of contrast modulated plaid patterns was dependent on the orientational bandwidth of the carrier and suggested that the extraction of second-order motion could not be accounted for by a “filter-rectify-filter” model.

To summarize, the bulk of the results in psychophysics suggest that, firstly, there is no fixed spatial-frequency relationship between the carrier and envelope, and secondly, that the processing of the carrier is probably orientation dependent (although its orientational relationship to the envelope remains unclear).

To examine the physiological mechanisms subserving second-order processing, we investigated the spatial characteristics underlying neurons’ responses to second-order envelope stimuli. We previously reported that the detection of envelope stimuli was dependent on both the carrier and envelope orientations (Mareschal & Baker, 1998a). In this paper, we extended this study by investigating the two-dimensional relationship between the carrier and envelope (spatial and orientational) as well as the bandwidth of the responses. Also, neurons’ responses to envelope stimuli were compared to their responses to luminance stimuli in order to shed light on similarities in their tuning to these different stimuli.

Indeed, it has been demonstrated in the macaque visual cortex that certain neurons respond in a similar manner to a stimulus irrespective of the physical cues defining it (“form-cue invariance”; Albright, 1992). The motivation of this study was to gain insight into the processing capabilities of individual neurons in the early stages of the visual system. A knowledge of these properties can then be used to guide both psychophysical and computational approaches to understanding second-order motion.

Methods

Animal preparation

Animal preparation was conventional and has been described in detail previously (Zhou & Baker, 1994). Experiments were carried out on 13 adult cats, paralyzed (gallamine triethiodide) and anesthetized (nitrous oxide/oxygen anesthesia supplemented with intravenous barbiturate). Vital signs (EEG, EKG, expired CO₂, and body temperature) were monitored and maintained at normal levels throughout the experiment. Platinum–iridium microelectrodes (Frederick Haer, Brunswick, ME) were used for penetrations in area 18, because of its higher proportion of envelope responsive neurons (Zhou & Baker, 1994). Each eye was refracted with a retinoscope and fitted with a gas-permeable neutral contact lens with an artificial pupil and spectacle lens in order to bring stimuli viewed at 57 cm into focus.

Stimuli

Two types of stimuli were used in these experiments: conventional sine-wave gratings and envelope stimuli, which consist of a stationary high spatial-frequency carrier whose contrast is modulated by a moving low-frequency envelope. The luminance profile of the envelope stimulus with independent carrier and envelope orientations is

\[
L_{\text{env}}(x, y, t) = L_c \left[ 1 + C \sin(2 \pi f_c (x \cos \alpha_c + y \sin \alpha_c)) \right] 
\]

\[
\times \left[ 1 + \sin 2 \pi (f_r (x \cos \alpha_r + y \sin \alpha_r) - f_t t) \right]/2, 
\]

where \(C\) is the carrier contrast which was set to 70%, \(f_r\) and \(\alpha_r\) are the carrier spatial frequency and orientation, respectively, and \(f_c, f_t, \) and \(\alpha_c\) the envelope spatial frequency, temporal frequency, and orientation.

Fig. 1 shows single frames of a drifting sine-wave grating (Fig. 1A) and an envelope stimulus (Fig. 1B), superimposed on a cartoon receptive field of a simple cell. In this example, the sine-wave grating is optimally matched to the neuron’s receptive field and should elicit a vigorous response (Fig. 1A). In the case of the envelope stimulus, the carrier spatial frequency is typically three octaves higher than the optimal frequency for the cell (Fig. 1B), and as such should equally stimulate inhibitory and excitatory subregions of the receptive field, leading to no response. In the frequency domain, idealized magnitude spectra of the stimuli are shown in Figs. 1C and 1D along with the luminance passband of the neuron (ovals). A sine-wave grating has two rotationally symmetric spectral components (Fig. 1C) where distance from the origin represents the spatial frequency and angular subtense represents the orientation. Fig. 1D shows the Fourier spectrum of the envelope stimulus, consisting of three spectral components (and symmetric counterparts in opposite quadrants) centered around the carrier frequency, with the two sidebands having frequencies equal
to that of the carrier plus and minus the envelope. There is no Fourier component at the envelope spatial frequency.

**Stimulus display and spike collection**

Previous studies have demonstrated (Zhou & Baker, 1993, 1996) that envelope responses were not the result of artifactual luminance nonlinearities arising from the screen. Some of their control experiments involved (1) measuring the effective contrast of tiny distortion products which could arise from the screen, (2) comparing responses to envelopes to responses to estimated distortion products, and (3) showing that optically low pass filtering the screen abolished envelope but not luminance grating responses. All luminance stimuli and the envelope stimuli with the carrier and envelope orientations locked (i.e. at the same orientations) were generated on the PC. However, for the carrier and envelope orientation experiments real time movies with independent carrier and envelope orientations could be generated only on the Macintosh.

**PC generated stimuli**

The stimuli were generated with a 66-MHz 80486 microcomputer, using a VSG 2/2 graphics card (Cambridge Research Systems, Cambridge, UK) and displayed on a NEC XP-17 monitor with a frame refresh rate of 160 Hz, a raster of 512 × 379 pixels, and a mean luminance of 28.6 cd/m². The luminance nonlinearity of the display was measured using a photometer (United Detector Technology, Orlando, FL, S-370), then linearized following the method of Pelli and Zhang (1991) using a modification of the VideoToolbox® software (Pelli, 1997) and an ISR Video Attenuator (Institute for Sensory Research, Syracuse University, NY) driving the green CRT gun. Stimulus motion was generated by lookup table (LUT) animation, which for the envelope stimuli consisted of digitally multiplying on each frame the profiles of the stationary carrier and the drifting envelope in the host computer. These were then used as indices to the Pelli-Zhang LUTs and downloaded to the graphics card LUTs in real time for each frame. Spike collection (RC-Electronics, Santa Barbara, CA, ISC-16) was computer controlled to an accuracy of 0.1 ms.

**Macintosh generated stimuli**

The stimuli were displayed on the same monitor with the frame refresh rate at 67 Hz, raster size of 640 × 480 pixels, and mean luminance 22.3 cd/m². The stimulus display was linearized by
gamma correction of the lookup tables. Envelope stimuli were produced as precalculated digital movies on a Power Macintosh using Matlab (The Mathworks, Natick, MA), VideoToolbox (Pelli, 1997) and Psychophysics Toolbox (Brainard, 1997) software. This setup was used exclusively for producing envelope stimuli with different orientations for the carrier and envelope components. Spike collection (Instrutech, Great Neck, NY, ITC-18) was computer controlled to an accuracy of 0.1 ms.

**Verification of the measurement compatibility using the two setups**

On a few representative neurons, we replicated certain quantitative measurements using stimuli produced by the PC and Macintosh setups. These tests included luminance frequency tuning and envelope carrier frequency tuning. In all cases, the optimal values of the varied parameters were not significantly different. In addition, other characteristics such as the orientation preference and direction selectivity were consistent for both methods of stimulus generation.

**Cellular recording**

Single-unit signals from area 18 were amplified and isolated with a window discriminator (Frederick Haer, Brunswick, ME), and monitored on a backward-triggered digital storage oscilloscope. Preliminary receptive-field mapping was done using a hand projector to determine the location, ocular dominance, preferred orientation, and eccentricity (average = 20 deg). Subsequently, the monitor was centered over the receptive field and computer-generated stimuli were presented to the neuron’s dominant eye. Test conditions lasted 3 s and were randomly interleaved, and spontaneous activity measured and removed from the subsequent poststimulus time histograms.

Drifting luminance-defined sine-wave gratings were used to measure the neuron’s tuning to spatial frequency, temporal frequency, and orientation. Initial testing with the envelope stimulus was done by setting the envelope spatial frequency and orientation to the neuron’s optimal luminance spatial frequency and orientation. The envelope temporal frequency was set slightly lower than the optimal luminance temporal frequency (typically 1 octave lower; Mareschal & Baker, 1998b), and a series of carrier frequencies much higher than the neuron’s luminance passband were tested (Zhou & Baker, 1993). Cells were classified as envelope responsive if their responses were direction selective, and showed an optimal carrier spatial frequency distinct from their tuning to luminance gratings (see Fig. 2, Mareschal & Baker, 1998b). Using the neuron’s measured optimal carrier spatial frequency, the envelope spatial frequency, temporal frequency, and orientation were varied independently. In this stage, the envelope orientation tuning was measured with the carrier and envelope orientations locked together. In subsequent tests using Macintosh stimuli, the envelope orientation was fixed at this measured optimum, and the carrier orientation was varied through 180 deg. Following this, the carrier orientation was set to its measured optimum and the envelope orientation was independently varied.

Poststimulus time histograms (PSTHs) were collected and integrated to obtain an average spike frequency as a function of the stimulus parameter being varied, for 33 envelope responsive neurons (out of a total of 73 neurons tested). Estimates of spatial-frequency bandwidth and optimum (for both the envelope and carrier) were obtained by fitting Gaussian functions to the response curves. Fig. 2 shows examples of the spatial-frequency tuning curves measured for one neuron. In Fig. 2A, spatial-frequency tuning was measured using luminance gratings and the data fitted with a Gaussian. Sample stimuli are shown below the tuning curve with spatial frequency increasing toward the right. In the envelope experiments, carrier spatial frequency was measured first with the envelope spatial frequency set slightly lower than the luminance optimum obtained in Fig. 2A. The carrier-frequency tuning curve (Fig. 2B) shows similar bandpass characteristics as the luminance tuning curve but for much higher frequencies. Illustrations of varying the carrier spatial frequency are shown below the tuning curve with examples of a low carrier spatial frequency in the left-hand plot, and a high carrier frequency in the right-hand plot. In Fig. 2C, envelope-spatial-frequency tuning is measured with the carrier spatial frequency fixed at the measured optimum. Illustrations of varying the envelope spatial frequency are shown below the graph (same format as above, with envelope spatial frequency increasing toward the right).

Once the optimal envelope spatio-temporal parameters were measured, orientation tuning was determined by plotting the normalized data in polar form where distance from the origin represents magnitude of response and angular subtense represents orientation. Estimates of bandwidth were obtained by calculating circular variance, given by the following equation (Marida, 1972; see also Ringach et al., 1997):

\[ V = 1 - \frac{\left| \sum_k R_k \exp(i2\theta_k) \right|^2}{\sum_k R_k^2} \]

where \( R_k \) is the strength of response for the given orientation \( \theta_k \) (in radians). Circular variance values range between 0 and 1, where 1 represents equal responsivity across all orientations (180 deg bandwidth), and 0 indicates responsivity to only one orientation. It is important to note that these bandwidth measurements are not equivalent to measurements of full-width at half-height, which would be obtained from data presented on a linear orientation axis. The advantage of a circular orientational index is two-fold; firstly, it provides orientational bandwidth indices for neurons having little or no orientation selectivity (for which full-width at half-height measures would be undefined). Secondly, since orientation is by definition a circular parameter, an index of selectivity obtained from data presented in circular form is a more natural measure.

Fig. 3 shows such polar plots of orientation responses from one neuron to three types of oriented stimuli, and presents circular variance indices which reflect the degree of selectivity in the different plots. Insets of the stimuli are shown alongside, depicting some of the different instances used. Fig. 3A shows the response to a sine-wave grating with the corresponding calculated circular variance (note the low circular variance index, which reflects relatively sharp tuning). In Fig. 3B, the envelope orientation is varied and the carrier orientation fixed at its measured optimum. In Fig. 3C, the carrier orientation is varied while the envelope orientation is fixed. Although the polar plot clearly shows orientation tuning, the value of circular variance is high due to the overall broader tuning to the carrier. (Note that in the experimental protocol, envelope orientation tuning was measured subsequent to determination of carrier tuning. In this figure it is shown before the carrier tuning to highlight the similarity between the envelope and luminance tuning curves.) The polar plot in Fig. 3C is symmetrical because the carrier was stationary; data were collected for the carrier orientation varying over
Fig. 2. Spatial-frequency tuning of an area 18 neuron to luminance and envelope stimuli. (A) Luminance spatial frequency tuning, examples of the stimuli are shown below the tuning curve (contrast = 30%, tf = 8 Hz). (B) Carrier spatial-frequency tuning, with the envelope fixed at the measured optimal spatial frequency. Examples of varying the carrier spatial frequency are shown below the tuning curve (contrast = 70%, tf = 4 Hz). (C) Envelope spatial frequency tuning, with a fixed carrier spatial frequency. Examples of stimuli are shown below the tuning curve. In our experiments, the envelopes were fully modulated although the illustration does not show this clearly.
180 deg and then reflected about the origin. The data presented in Fig. 3A were obtained using stimuli generated on the PC, data in Figs. 3B and 3C were obtained with Macintosh-generated stimuli. Statistical comparisons between measurements on populations of

Fig. 3. Orientation and direction tuning of an area 18 neuron to luminance and envelope stimuli. (A) Normalized responses as a function of the orientation of a sine-wave grating (lum sf = 0.132 cpd, tf = 4 Hz) (three possible configurations shown alongside). (B) Responses from the same neuron as a function of the envelope orientation with a fixed carrier orientation (env sf = 0.093 cpd, env tf = 3, carr sf = 1.5 cpd). (C) Carrier orientation tuning when the envelope orientation is fixed near the optimum.

Fig. 4. Optimal spatial frequencies to envelope and luminance stimuli. (A) Each point represents one neuron's optimal luminance spatial frequency plotted against its optimal envelope spatial frequency. Solid lines depict the unity and 2:1 ratios. (B) Same neurons as in A, but plotting the optimal carrier spatial frequency against the optimal luminance spatial frequency. Solid lines depict fixed ratio relationships. (C) Same neurons as in B, but plotting the optimal carrier spatial frequency against the optimal envelope spatial frequency.
Results

Spatial scale

Neurons’ relative spatial tuning (peak and bandwidth) to the carrier and envelope was measured in order to quantitatively compare the two kinds of spatial selectivity. Peak spatial frequencies and bandwidths were measured in 33 neurons using luminance gratings and envelope stimuli presented at the optimal temporal frequencies (average optimal luminance temporal frequency = 5.23 Hz, average optimal envelope temporal frequency = 2.88 Hz). This is plotted in Fig. 4 where each point represents the optimum obtained from the Gaussian fit for one neuron to the different stimuli. Fig. 4A plots the optimal luminance spatial frequency (avg = 0.08 cpd) against the optimal spatial frequency to the envelope (avg = 0.066 cpd) (with the carrier fixed at the measured optimum). Note that many points fall above the 1:1 ratio line; on average neurons are tuned to significantly lower envelope spatial frequencies. There is little correlation between these two parameters (r = 0.4).

In Fig. 4B, the optimal carrier spatial frequency (avg = 1.09 cpd) is plotted against the optimal luminance spatial frequency. In these experiments some of the carrier spatial frequencies approach the cat’s acuity limit (roughly 7 cpd in central vision, Ulrich et al., 1981; see also Zhou & Baker, 1996, Fig. 6). On average, the optimal carrier spatial frequency was approximately 10-fold higher than the optimal luminance spatial frequency. This was not, however, a fixed relationship; ratios ranged from 5- to 30-fold with very little correlation (r = 0.2).

Fig. 4C plots the optimal carrier spatial frequency against the optimal envelope spatial frequency and reveals a slightly more pronounced spatial-frequency difference (on average, 15-fold higher for the carrier). As for the ratios in Fig. 4B, the values were not fixed and ranged from 8- to 33-fold (r = 0.3). The dashed line represents the minimum spatial separation between the carrier and envelope that we tested. Indeed, as the carrier and envelope spatial frequencies approach one another, luminance information in the stimulus (the side-band information) could approach the neuron’s luminance passband. To avoid any possible such luminance-based artifacts, we chose a minimum separation of 3 octaves between the carrier and envelope spatial frequencies.

Spatial frequency bandwidths for 22 neurons (full-width at half-height) were calculated from the Gaussian fits and plotted as histograms (Fig. 5). We previously found that there was not a significant difference in the spatial bandwidths measured to the luminance gratings and to the envelope modulation (in area 18, lum avg = 1.43 octaves; Mareschal & Baker, 1998b). Here we compare the spatial bandwidths for the carrier and envelope tuning. Fig. 5A plots the envelope spatial-frequency bandwidths (avg = 1.46 octave) whose values are comparable to values measured in area 17 using luminance gratings (normalized range between 0.3 to 3 octaves, avg = 1 octave; Tolhurst & Thompson, 1981). The carrier bandwidth values are narrower than the bandwidth values measured to luminance gratings or the envelope, but are still within the range of values previously reported in area 17 (Fig. 5B, avg = 0.95). Envelope and carrier bandwidth values are plotted against one another in Fig. 5C where the solid line represents equality. We found that the carrier bandwidths are significantly narrower than the envelope bandwidths (two-sided Students t-test, P < 0.05) and not correlated.

Fig. 5. Spatial-frequency bandwidths measured as full-width at half-height. (A) Distribution of envelope spatial bandwidths. (B) Distribution of carrier spatial-frequency bandwidths. (C) Envelope spatial-frequency bandwidths plotted against carrier spatial bandwidths. Solid line shows equity relationship.
Orientation tuning of envelope-responsive neurons, quantified as peak and bandwidth parameters, was measured for luminance and envelope stimuli. Results from three representative neurons are shown in polar plots in Fig. 6 (same format as Fig. 3). Fig. 6A (top row) shows orientation tuning measured with luminance gratings, and demonstrates that these neurons are narrowly tuned and often strongly direction selective to gratings. In Fig. 6B, responses are plotted for the same neurons using envelope stimuli with the carrier orientation held constant and the tested values of envelope orientation spanning 180° for both directions of motion. The data resemble the orientation responses measured with the gratings: similar peak orientations, tuning curves (although slightly higher values of circular variance, with the exception of F3506) and amount of direction selectivity. Responses measured when the carrier orientation was varied, while keeping envelope orientation fixed, are plotted in Fig. 6C (note the carrier was stationary so symmetric data points are reflected about the origin). A number of important differences emerge from these plots. First, the carrier responses were clearly orientation selective, although always more broadly so that the responses obtained to the envelope or the sine-wave grating. Also, the responses rarely fell to zero at nonoptimal orientations, resulting in larger values of circular variance. Finally, there does not appear to be a fixed relationship between the optimal carrier and envelope orientations. For example, the optimal envelope and carrier orientations for cell F3506 appear to be only slightly different, while for cell F3703 they appear to be nearly orthogonal.

In some neurons, the optimal carrier or envelope orientation (when these were measured independently) differed from that measured initially when the envelope and carrier orientations were locked. To assess whether changing the orientation of the parameter being held constant (carrier or envelope) had an effect on the measured peak and spread of the response to the other parameter, we measured in one neuron envelope tuning for two different carrier orientations (Fig. 7A). The top polar plot is the envelope orientation tuning obtained when the carrier orientation was set to its measured optimum. The bottom plot was measured when the carrier orientation was set to the optimum obtained when the envelope and carrier orientations were varied together. In this example, the two different carrier orientations differ by 45°. The overall form of the response was similar in the two conditions, although the absolute strength of the response was reduced (the strength of the peak response was 72% that of the top peak response). In Figs. 7B and 7C, we measured the carrier orientational response for two different fixed envelope orientations. For both neurons, the top plots were obtained when the envelope orientation was set to the optimum measured when the envelope and carrier orientations were varied together. The bottom plots were measured when the envelope was set to its measured optimum (obtained independently of the carrier). In Fig. 7B, the difference in envelope orientations between the top and bottom conditions was 25°, in Fig. 7C it was 67.5°. The carrier orientation response in Fig. 7B was comparable in the two conditions (strength of the response did not vary). In Fig. 7C, the response shifted slightly from the optimum obtained in Fig. 7B and was reduced by 40%. However, in this condition, the orientation of the envelope is nearly perpendicular to its measured optimal orientation (used in the top portion of Fig. 7C). Arrows on the lower plots indicate where the peak would be expected if the optimal carrier and envelope orientations were locked (e.g. shifting the peak of one by a certain amount would...
shift the peak of the other by the equivalent amount). This clearly does not appear to occur in neurons’ orientation tuning, suggesting that the protocol of measurements used here had a minimal impact on the results. That is to say, our estimate of one parameter’s optimal orientation (either the carrier or the envelope) was relatively independent of the orientation of the other parameter.

To assess the relative degree of orientation selectivity to luminance gratings and envelope stimuli, each neuron’s circular variance values to these different parameters was directly compared. In Fig. 8A, each point represents one neuron’s measured circular variance index to a luminance grating against its circular variance index to the envelope. Indices of circular variance are directly related to bandwidth where a value of 1 indicates no tuning (isotropic response) and 0 reflects sharp tuning. Most points lie below the unity ratio line, revealing that in general, neurons’ orientation tuning to the envelope (avg = 0.48) is broader than to a luminance grating (avg = 0.26). Fig. 8B plots circular variance indices measured to the envelope versus the carrier (avg = 0.84). Note that there is no correlation between the envelope and carrier orientation tuning. Also, circular variance indices are significantly higher for the carrier than the envelope, due at least partially to the responses rarely falling to zero at nonoptimal carrier orientations (see, e.g. cell F4002 in Fig. 5C). A possibly more revealing index would be to estimate the degree of orientation selectivity between the maximum and minimum values in the orientation plots. This was done by subtracting the lowest firing rate value in the orientation plot from all other values prior to calculating circular variance. Fig. 8C plots the same data as Fig. 8B, but using these “normalized” carrier values. It is apparent that even after normalization the circular variance indices are quite high (avg = 0.6), indicating that orientation bandwidths for the carrier are quite broad.

In Figs. 3 and 6, it is apparent that the optimal envelope, carrier, and perhaps luminance orientations can differ within the same neuron. To evaluate whether any of these orientations were systematically related (e.g. whether on average the orientation difference between the optimal carrier and envelope tended to be at some fixed value), we compared peak orientations in each neuron measured as a function of the different parameters. Each point in Fig. 9A represents a neuron’s measured optimal orientation of luminance grating against its optimal orientation of envelope for the preferred direction of motion. The solid line represents the equality ratio (same optimal orientations), whereas the two dashed lines represent orthogonal preferred orientations. Most points cluster around the equality line, except for two neurons whose optimal orientations differed by more than 45 deg. The optimal orientations measured to these two parameters were quite strongly correlated (r = 0.85) suggesting that the neuron’s response was in this sense “form-cue invariant.” That is, a neuron’s selectivity is the same regardless of the physical cue defining the stimulus (here luminance vs. contrast). When the optimal orientation to the carrier was plotted against the optimal envelope orientation (Fig. 9B), there was no fixed orientational relationship between them (r = 0.25). To search for an underlying relationship between the carrier and envelope orientations which might not emerge from a scatterplot (e.g. if there tended to be two or three orientational relationships between the carrier and envelope), we plotted the absolute value of the difference between the carrier and envelope orientations (Fig. 9C, unfilled histogram). Psychophysical results suggest that parallel and perpendicular relative orientations between the carrier and envelope are more easily detected (Dakin & Mareschal, 1998), so we chose large binwidths to examine this. While all neurons had similar orientations to the luminance and envelope (filled histogram, unimodal distribution), there was no apparent bimodal distribution in the optimal orientations between the envelope and carrier.
Discussion

Examination of neurons’ tuning to the carrier component of the envelope stimulus revealed two defining characteristics: that the carrier must be of a much higher spatial frequency than the envelope, and that the neurons which process the carrier are orientation selective. The clear dependence of neurons’ responses on both the orientation and spatial frequency of the carrier clearly discard an early nonlinearity as an explanation for neuronal responses to envelope stimuli. These results provide further support for two parallel streams of motion processing, one a linear stream processing luminance-defined stimuli, the other a nonlinear stream processing the contrast-defined stimuli.

Nonlinear stream: Carrier processing

Previous physiological (Zhou & Baker, 1993, 1994, 1996; Mareschal & Baker, 1998b) and psychophysical (Sutter et al., 1995) studies reported selectivity for high carrier spatial frequencies. However, the relationship between carrier and envelope spatial frequency had only been addressed psychophysically, and the results were contentious. The bulk of the findings suggested no fixed carrier/envelope spatial frequency relationship (Jamar & Koenderink, 1985; Cropper, 1998; Dakin & Mareschal, 1998). However, one study implied an optimal carrier/envelope spatial-frequency ratio (Sutter et al., 1995), and another suggested a fixed optimal envelope spatial frequency independent of the carrier spatial frequency (Langley et al., 1996). Here, we report that although optimal carrier frequencies are on average 15-fold higher than for the envelope, there is no fixed relationship between them.

Orientation tuning to the carrier has only been examined psychophysically and the bulk of the results suggest oriented carrier processing. Wilson et al. (1992) found that the processing of second-order motion components in a plaid pattern required spatial-frequency and orientation-selective filtering. Using an adaptation paradigm, Langley et al. (1996) report that the processing of the carrier is spatial frequency and orientation selective. Cropper (1998) and Dakin and Mareschal (1998) also suggest the use of bandpass-oriented filters for the processing of the carrier. However, Werkhoven et al. (1994) and Boulton and Baker (1994) both suggested the use of broadband spatial and orientational carrier processors (termed “texture-grabbers” by Werkhoven et al. 1994). Our results clearly support the notion of orientation-dependent filtering of the carrier.

Based on the orientation tuning results, we believe that the neurons processing the carrier are cortical, despite evidence for orientation tuning in some geniculate neurons (Vidyasagar & Heide, 1984; Soodak et al., 1987; Shou & Leventhal, 1989; Smith et al., 1990) and claustral neurons (Sherk & LeVay, 1981). The above experiments reported that certain LGN neurons displayed orientation tuning when tested with gratings whose spatial frequency was much higher than a given neuron’s measured optimum. However, when the same LGN neurons were tested with gratings at their

Fig. 8. Circular variance values to the envelope and luminance stimuli. (A) Luminance circular variance indices plotted against envelope circular variance indices (correlation, $r = 0.66$). Solid line depicts equality. (B) Envelope circular variance indices plotted against carrier circular variance indices ($r = 0.04$). (C) Envelope circular variance indices plotted against “normalized” carrier circular variance indices ($r = 0.14$).
optimal spatial frequency, they displayed very little orientation selectivity. This was mainly interpreted as the ganglion cell center mechanism being made up of smaller aligned subfields which would be stimulated by high spatial-frequency gratings, leading to orientation biases. In our experiments, the carrier orientation tuning was measured using the cortical neuron’s optimal carrier spatial frequency, which was quite high relative to its optimal luminance spatial frequency. Thus one might suppose that the orientation tuning of LGN cells at high spatial frequencies might provide a source of cortical neurons’ orientation tuning. This explanation, however, requires that the geniculate inputs onto the second-stage neuron have consistent orientational biases and be aligned. Although not impossible, such an organization seems highly unlikely.

Claustral neurons display quite narrow orientation tuning and could conceivably provide the basis for processing of the carrier (Sherk & LeVay, 1981; LeVay & Sherk, 1981). However, the orientation selectivity in the claustral neurons is believed to be mediated via feedback connections from the cortex (areas 17 and 18) onto these neurons. In this sense, even if the neurons which process the carrier were claustral, the origin of their orientation selectivity would be of a cortical nature.

Many neural wiring schemes could produce the broad carrier orientation bandwidths observed here. Broad orientation tuning might be the direct result of broadly oriented, high spatial-frequency cortical neurons. There is evidence in the monkey that high spatial-frequency layer 4C cells display broad orientation tuning (Ringach et al., 1997). If a similar organization existed in cat area 17, these neurons could provide the basis for the carrier processing. Alternatively, the broad orientation tuning could result from the combined inputs of many narrowly tuned, slightly differently oriented, high spatial-frequency neurons. In either case, area 17 neurons have been found which are tuned to high spatial frequencies (values ranged from <0.3 cpd to >3 cpd) similar to those of the carrier tuning (Tolhurst & Thompson, 1981).

Nonlinear stream: Envelope processing

In contrast to the results for the carrier tuning, the selectivity of cortical neurons for the envelope was more similar to the luminance grating results. Firstly, although the spatial and orientational bandwidths are broader for the envelope, these differences are not very compelling (particularly compared to the carrier and envelope differences). Secondly, the optimal orientations to the envelope and the sine-wave grating are almost always very similar. We previously reported that the optimal spatial and temporal frequencies were significantly lower to envelope than to luminance gratings (Mareschal & Baker, 1998). The emerging picture from these comparisons is that the processing of the envelope involves neurons tuned to lower spatial frequencies with slightly broader

Fig. 9. Measured optimal orientations to first- and second-order stimuli. (A) Results for 24 neurons where each point represents one neuron’s measured optimal orientation to a sine-wave grating (ordinate) against its optimal orientation to the envelope (abscissa). The dashed lines represent orthogonal optimal orientations between the abscissa and ordinate. (B) Results for 22 of the neurons in A, but plotting the optimal orientation to the carrier (ordinate) against the optimal orientation to the envelope (abscissa). (C) Histogram of orientation differences between envelope and a luminance grating (solid) and between the envelope and the carrier (open).
Possible factors affecting our results

Although we find that neurons are tuned to much higher frequencies for the carrier than the envelope, it is important to note that this may be a direct result of our protocol to ensure that we are effectively measuring “second-order responses.” If the carrier and envelope spatial frequencies are too close together, two possible artifacts can occur.

In the first instance, as the spatial frequency of the carrier approaches that of the envelope, the energy in the stimulus (particularly in the lower sideband) may fall within the neuron’s luminance-defined passband. For illustrative purposes, in Fig. 1D the three components are centered around the carrier spatial frequency. By testing lower carrier frequencies, these components are shifted toward the origin (lower spatial frequencies), where the response elicited by the envelope stimulus could simply result from the neuron’s response to luminance components (i.e. a linear response). We specifically screened for neurons with a large spatial-frequency difference between the carrier and envelope to avoid this artifact. When this constraint is relaxed, cortical neurons have been found to respond to envelope stimuli with closer envelope and carrier spatial frequencies (O’Keefe & Movshon, 1998), but it is unclear whether such responses are “second-order.” We do not refute the existence of carrier processing neurons tuned to low spatial frequencies, but our method of assessment does not allow us to disambiguate such responses from simple linear stream responses.

The second issue arises from the fact that varying the relative orientations of the carrier and envelope can introduce a small skew in the orientation response function, which in turn would cause small changes in the estimated peak orientation and bandwidth. Because our experiments were conducted with a 3-octave minimum separation between the carrier and envelope spatial frequencies any subtle differences in the frequency composition of the stimulus will be centered around the carrier frequency, which is constrained to be far outside the neuron’s luminance passband. We assessed the extent of the skew on estimated measurements by convolving envelope stimuli with carrier-tuned filters which were oriented (one-dimensional Gabor functions) or isotropic (Difference of Gaussian) filters. In the case of the oriented filters, we found that there could be a small influence of the relative orientations of the stimulus; however, the exact magnitude of this effect depends on both the filter parameters and the contrast response function, and cannot be quantified independently. In the case of simple, linear, isotropic filters we found that varying the relative carrier and envelope orientations was insufficient to induce any significant orientational biases in these filters (estimates of circular variance from the filtering were 0.99). The stimulus configuration could not make an isotropic first-stage filter appear significantly oriented.

Implications for models of second-order motion processing

The detection of second-order stimuli is most often accounted for by a model of motion processing consisting of parallel linear and nonlinear streams (Graham et al., 1992; Wilson et al., 1992; Zhou & Baker, 1993; Nishida et al., 1997; but see Johnston et al., 1992). In this model, stimuli defined by changes in luminance are processed by the linear stream and stimuli defined by changes in contrast are processed by the nonlinear stream. In the nonlinear stream, the motion or orientation signal of a contrast-modulated (second-order) stimulus is made available to spatiotemporal filters via a nonlinear transformation after an initial bandpass filtering stage. It is essential to note that a sine-wave grating at the spatial frequency of the carrier would not elicit a response from the nonlinear stream, because its rectified representation would fail to elicit a response from the second filtering stage which is tuned to low spatial frequencies.

Fig. 10 illustrates alternative types of two-stream motion models where the nonlinear stream consists of a pool of first-stage neurons tuned to higher spatial frequencies than the second-stage neurons. In Model 1 the processing of the carrier involves isotropic filters, presumably located in the lateral geniculate nucleus, whereas Model 2 posits oriented filters located in the cortex. Our results
support Model 2, and provide certain insights into the architecture of this cortical-based model. Although the first-stage filters presumably tile the orientation space, their inputs onto the second-stage filters would be weighted in an orientation-dependent manner. This would provide broad orientation tuning to the carrier, with an orientational peak corresponding to the strongest weighted filter. The tuning of the second-stage filters resembles that of the filters in the linear stream, but is slightly broader (in frequency and orientation) and tuned to lower spatial and temporal frequencies.

The lack of systematic relationship between the carrier and envelope orientations could be functionally important, because it should endow these neurons with the capability of responding to a broader variety of second-order stimuli. For example, neurons have been reported which respond to “illusory contour” stimuli defined solely by changes in phase (Von Der Heydt et al., 1984; Grossoft et al., 1993; Sheth et al., 1996). A processing scheme of the type depicted in Model 2 might account for the detection of such stimuli without requiring a separate, additional, nonlinear processing stream.

Acknowledgments

We thank Steven Dakin for providing valuable comments on the preparation of this manuscript, David Badeck for insightful discussion of the stimuli, and an anonymous reviewer for helpful suggestions on Fig. 2. We are also grateful to Jingjiang Lei and Lynda Domazet for technical assistance, and Ken Charles and Michael Moskovich for computer programming. We wish to thank Rhone-Poulenc Rorer for their generous donation of Gallamine triethiodide. This work was supported by a Canadian MRC grant (MA 9685) to C.L. Baker and an FCAR fellowship to I. Mareschal.

References


