Title: Motion-defined contour processing in early visual cortex.

Authors: AMOL GHARAT and CURTIS L. BAKER Jr.

Author Contributions: This study was designed by both Amol Gharat and Curtis Baker. Animal preparation for the recording experiments was done by Amol Gharat and Curtis Baker, with assistance from other lab members. The software for visual stimulus presentation and data acquisition was primarily built by Curtis Baker and was further modified by Amol Gharat. Most of the data was collected by Amol Gharat. All the data analysis and writing of the manuscript was performed by Amol Gharat with guidance from Curtis Baker.

Affiliation: Department of Psychology, McGill University; McGill Vision Research Unit, Department of Ophthalmology, McGill University, Montreal, Quebec, Canada.

Running head: Motion-defined contour processing.

Contact information: Amol Gharat, Department of Psychology, McGill University, 1205 Dr. Penfield Avenue, Montreal, Quebec, H3A 1B1, Canada. Tel. no. +1 514 815 3510. E-mail: amol.gharat@mail.mcgill.ca
Abstract

From our daily experience it is very clear that relative motion cues can contribute to correctly identifying object boundaries and perceiving depth. Motion-defined contours are not only generated by the motion of objects in a scene, but also by the movement of an observer’s head and body (motion parallax). However the neural mechanism involved in detecting these contours is still unknown. To explore this mechanism, we extracellularly recorded visual responses of Area 18 neurons in anesthetized and paralyzed cats. The goal of this study was to determine if motion-defined contours could be detected by neurons that have been previously shown to detect luminance, texture- and contrast-defined contours cue-invariantly. Motion-defined contour stimuli were generated by modulating the velocity of high spatial frequency sinusoidal luminance gratings (carrier gratings) by a moving squarewave envelope. The carrier gratings were outside the luminance passband of a neuron, such that presence of the carrier alone within the receptive field did not elicit a response. Most neurons that responded to contrast-defined contours also responded to motion-defined contours. The orientation and direction selectivity of these neurons for motion-defined contours was similar to that of luminance gratings. A given neuron also exhibited similar selectivity for the spatial frequency of the carrier gratings of contrast- and motion-defined contours. These results suggest that different second-order contours are detected in a form-cue invariant manner, through a common neural mechanism in Area 18.

Keywords

Figure-ground segregation, Second-order motion, Relative motion.
Introduction

Natural scenes abundantly contain local variations in luminance that facilitate figure-ground segregation. However, these first-order cues often introduce ambiguities and make figure-ground segregation a difficult task (Marr 1982). For example, shadows introduce false luminance boundaries that do not correspond to objects’ boundaries in a visual scene. However, our visual system is able to distinguish these false boundaries from real ones using other cues, including second-order information such as texture, contrast, color, or motion differences between an object and its background. Particularly, relative motion is a powerful cue that can break camouflage when an object and its background have similar luminance, color, and texture. It can be sufficient to support perception of shape and size of three-dimensional surfaces, and for depth ordering (Rogers and Graham 1979; Regan 1989; Regan and Hamstra 1992). This cue arises from motion parallax generated by an observer’s movement, or from exogenous movement of objects in a scene.

Even though psychophysical studies have demonstrated the importance of relative motion cues in figure-ground segregation, the neural mechanism to detect these motion-defined boundaries is still unknown. Single-unit recording experiments by (Hubel and Wiesel 1962) on cats showed that orientation selectivity for luminance edges first originates in brain areas as early as primary visual cortex (V1). Simple cells in V1 have receptive fields with elongated excitatory and inhibitory areas lying adjacent and parallel to one another, which act as filters that perform linear summation of light intensity in their receptive fields. (Hubel and Wiesel 1962) proposed a model in which receptive fields of simple cells are constructed by inputs from on-centre and off-centre LGN cells arranged in alternating columns. A similar question could be asked in the case
of motion-defined boundaries, i.e. where does orientation selectivity for these boundaries
originate and what is the neural mechanism behind it?

Several single-unit studies have tried to locate the brain areas responsive to motion-defined boundaries and understand the underlying neural mechanism. Using temporal texture bars (dynamic random dot patterns moving on a stationary random dot background), (Albright, 1992) reported that most of the neurons in area MT of macaque monkeys were selective for orientation of these bars and (Chaudhuri et al. 1997) found more than half of the neurons in area V1 of macaque monkeys selective for orientation of these bars. (Marcar et al. 2000) also found a small fraction of neurons in macaque Areas V1 and V2 that were selective for orientations of motion-defined boundaries. In macaque Area V4, (Mysore et al. 2006) reported a sizeable fraction of neurons (10-20%) that were selective for kinetic patterns. Both these studies (Marcar et al. 2000) and (Mysore et al. 2006) used moving random dot texture patterns to generate motion-defined boundaries which were held stationary in the receptive field of a neuron. (Zeki et al. 2003) found a majority of neurons in macaque Area V3 and V3A selective to orientation of motion-defined bars made of random dot texture patterns. These studies suggest that motion-defined boundary selective neurons are present in different visual areas such as V1, V2, V3, V3A, V4, and MT with higher cortical areas containing a greater percentage of such cells.

However there is a potential problem with the random dot texture patterns used in all these studies, because such textures contain a broad range of spatial frequencies. Hence these texture patterns will contain energy within the luminance passband of a neuron, and therefore the response of a neuron could be due to local luminance (first-order) signals and not motion difference (second-order) cues. Such luminance signals or artifacts can be avoided by using a
sinusoidal grating as a texture pattern, with the spatial frequency higher than the neuron’s luminance resolution.

A neuroimaging study in human subjects (Reppas et al. 1997) found strong motion boundary selective signals in early cortical areas (V1, V2). But it is unclear from neuroimaging studies whether the neurons in brain areas that respond to motion boundary stimuli are selective for the orientation of these boundaries, as neurons could be responding due to a centre-surround antagonistic mechanism (Born & Tootell 1992; Born 2000; Shen et al. 2007) or just to the local motion of the carrier. However, a recent study (Larsson et al. 2010) was able to demonstrate orientation selectivity to motion boundaries in human visual cortex using an event-related fMRI adaptation technique. They showed that most of the motion boundary responsive visual areas like V2, V3, V3A, V3B, LO1, LO2, hV4 and V7 identified in previous neuroimaging studies (Dupont et al. 1997; Larsson and Heeger 2006; Tyler et al. 2006; Van Oostende et al. 1997; Zeki et al. 2003) are orientation selective. These results argue against the initial notion from neuroimaging studies (Dupont et al. 1997; Van Oostende et al. 1997) that motion boundaries are processed in a specialized brain area, “kinetic occipital” or KO (corresponding to LO1, LO2, and V3B).

Neuronal responses to contrast-defined (second-order) boundaries have been extensively studied in cat Area 18 using single-unit recordings (Zhou & Baker, 1993, 1996; Mareschal and Baker, 1998a, 1998a, 1999). Contrast-defined boundaries used in these studies were constructed by a coarse spatial scale contrast pattern (envelope), which modulates the contrast of a high spatial frequency sinusoidal grating (carrier) (Fig. 1 B). Around half of the neurons in Area 18 responded to contrast-defined boundaries in a form-cue invariant manner, i.e. they were tuned to the same orientation and motion direction of luminance (Fig. 1 A) (first-order) and contrast-
defined (second-order) boundaries. In these studies carrier spatial frequency was constrained to lie outside a neuron’s spatial frequency passband (measured using luminance grating) to ensure that the response of a neuron was genuinely second-order and not due to first-order luminance signals. Surprisingly, these neurons showed narrow band-pass tuning for carrier spatial frequency. (Song and Baker 2007) subsequently showed that these contrast-defined boundary responsive neurons also respond to texture-defined boundaries (second-order) and again in a form-cue invariant manner. Texture-defined boundaries (illusory contours), similar to contrast-defined boundaries, were constructed using high spatial frequency sinusoidal gratings as a carrier, whose phase was modulated by a square wave envelope. Neurons showed narrow band-pass tuning for carrier spatial frequency of texture-defined boundaries and were selective for similar carrier spatial frequencies. These results suggest that these neurons would be functionally useful in mediating responses to boundaries regardless of the cue that defines them, and this cue-invariance to different second-order boundaries might arise from a common nonlinear neuronal mechanism.

First- and second-order information are thought to be processed by two parallel pathways, based on results from both psychophysics (Ledgeway and Smith 1994; Mather and West 1993; Nishida et al. 1997; Scott-Samuel and Georgeson 1999; Allard and Faubert 2007) and neurophysiology (Zhou and Baker 1993) - for review, see (Baker and Mareschal 2001). First-order information can be detected by neurons acting as quasi-linear spatio-temporal filters. For detecting second-order information, neurophysiology experiments support a two-stage filter-rectify-filter (FRF) model (Mareschal and Baker 1999), involving early filtering which is selective for local texture characteristics, followed by rectification, and a second-stage coarse-scale spatio-temporal direction selective filtering (Chubb & Sperling, 1988; Wilson 1999; Landy
& Graham 2004). The second stage filter has similar properties as the first-order filter, but it pools across a coarser spatial scale. These two parallel pathways converge onto a single neuron. This model is supported by recent optical imaging (Zhan and Baker 2006) and single-unit neurophysiology (Song and Baker 2007).

In this study we hypothesize that these second-order responsive neurons in Area 18 might also respond selectively to motion-defined boundaries, and that they might do so in a form-cue invariant manner. We used high spatial frequency sinusoidal gratings as texture (carrier) patterns, and relative motion between these textures to create motion-defined boundaries (Fig. 1 C & D). To ensure that responses were not simply due to the carrier motion, we carefully optimized the spatial frequency of the carrier grating for each neuron, such that it was well outside of the neuron’s conventional luminance grating resolution. A common motion-defined boundary occurs when an object moves in the visual field. In this case the retinal image of the background is nearly stationary, but the image of the object moves – we mimic this situation with a square wave envelope in which alternate half cycles contain either a moving or a stationary texture (carrier) (Fig. 1C) – a “uni-directional” motion boundary. We mimic motion boundaries generated from motion parallax with a stimulus in which texture in alternate half-cycles of the envelope moves in opposite directions (Fig. 1D) - a “bi-directional” motion boundary. We restricted this study to “shear” motion boundaries, in which local motions are parallel to the edge, to avoid complexities of accretion-deletion cues (Sary et al. 1994). To assess form-cue invariance, we compared neurons’ responses to motion-defined boundaries with those to contrast- and luminance-defined boundaries. For all three types of boundaries, the envelope was drifting at a low temporal frequency. These comparisons also enabled inferences regarding similarity between underlying neural mechanisms for these different stimuli. In addition, we
simulated a model of a generic Area 18 neuron receiving inputs from two parallel pathways that separately process first- and second-order information as described earlier, in order to see whether the selectivity of neurons to contrast- and motion-defined boundaries can be explained by a single such model.

We found that all contrast-defined boundary-responsive neurons also respond to uni-directional motion boundaries in a form-cue invariant manner, and with similar carrier spatial frequency tuning. Some but not all contrast-defined boundary-responsive neurons also respond to bi-directional motion boundaries, typically with weaker responses than to uni-directional boundaries. The pattern of selectivity of these neurons matches well with the selectivity of the simulated model. This suggests that motion-defined boundaries are processed by the same nonlinear neural mechanism that processes contrast-defined boundaries.
Materials and methods

Animal preparation

All experimental procedures were reviewed and approved by the Animal Care Committee of McGill University. Cats were anesthetized using isoflurane/oxygen, and maintained with isoflurane inhalation. Methylcellulose gel (1%) was applied to protect the corneas, and a rectal thermistor inserted to monitor temperature during surgery. Intravenous cannulation was performed and a loading dose of propofol (5 mg/kg) was delivered, and then maintained at 6 mg/kg/hr for subsequent surgery. EKG leads were connected to monitor heart rate. Tracheal intubation was performed to provide a secure airway and the animal was secured on a stereotaxic apparatus. A respirator (Ugo Basile) was connected to deliver a mixture of O2/N2O (30:70). End-tidal CO2 was monitored with a capnometer (Hewlett-Packard) and maintained between 28 to 36 mmHg by adjusting the respirator stroke volume. A pulse-oximeter sensor (Nonin) measured blood oxygen. Eye drops (atrophine 1% and phenylephrine 2.5%) were applied, and neutral contact lenses inserted. A craniotomy was made to expose Area 18 (Horseley-Clarke A3/L4, Tusa et al. 1979), and also a duratomy when recording with multi-electrodes. The craniotomy was covered with 2% agarose, followed by petroleum jelly. All surgical wounds were infused with bupivacaine (0.5 %) and the temperature was thermostatically regulated (Harvard Apparatus) at 37.5 °C. The animal was anesthetized and paralyzed with a continuous infusion of propofol (5.3 mg/kg/hr), fentanyl (7.4 μg/kg/hr) and gallamine triethiodide (10mg/kg/hr). Glycopyrrolate (0.005 mg/kg) and dexamethasone (0.6 mg) were delivered intramuscularly every 12 hours throughout the experiment. Artificial pupils were positioned, and appropriate spectacle lenses were selected using a slit retinoscope to provide refraction at a
viewing distance of 57 cm. The optic disk was back-projected on a tangent screen (Fernald and Chase 1971), and used to estimate the location of the area centralis of each eye (Nikara et al. 1968).

**Visual stimuli**

Visual stimuli were generated by a Macintosh (Intel 4x2.66 GHz, 6GB, NVIDIA GeForce GT 120) using Matlab with Psychophysics Toolbox (Brainard 1997; Pelli 1997) and presented on a gamma-corrected 17-inch CRT monitor (resolution 640x480 pixels, 75Hz). The stimuli were confined within 480x480 pixels corresponding to 30x30 deg at a viewing distance of 57 cm. Conventional luminance modulation (LM) sine wave gratings (Fig. 1 A) with a contrast of 30% were used to measure the luminance-passband of a neuron (spatial frequency, orientation and temporal frequency tuning). Contrast modulation (CM) gratings (Fig. 1B) were constructed by modulating the contrast of a carrier (texture pattern) by a low spatial frequency grating of 100% modulation depth (“envelope”). The CM grating was drifted with a temporal frequency slightly lower than the neuron’s optimum for LM gratings (Mareschal and Baker 1999). A high spatial frequency sinusoidal grating was used as a carrier, with a contrast of 70%. This carrier grating was held stationary, except for measurements of carrier temporal frequency selectivity in which it was drifted with varying temporal frequency. Motion-defined boundaries were generated using “velocity modulation” (VM) gratings in which alternate half cycles of the envelope contained a texture (carrier) moving with different velocities. This envelope was parallel to the motion direction of the carrier (“shear”), and it drifted in a direction perpendicular to the carrier motion with the same temporal frequency (between 1 to 4 Hz ) used for CM gratings. In particular, we tested two types of velocity modulation gratings, viz. uni-directional (Fig. 1 C) and bi-directional (Fig. 1 D). In uni-directional VM gratings, alternate half cycles of
envelope contained moving or stationary carrier. Bi-directional boundaries were created by oppositely moving carriers. For CM gratings, the envelope was sinusoidal, while for VM gratings it was a square-wave. All stimuli were presented within a raised cosine-tapered, circular window against a gray background of the same mean luminance. On some trials a uniform gray screen was presented to measure spontaneous activity.

**Extracellular recording**

Spikes from single neurons were recorded extracellularly with glass-coated Platinum/Iridium and parylene-coated tungsten single channel microelectrodes (Frederick Haer), and 16 channel multielectrodes (Neuronexus). Spike times were collected through a lab interface (Instrutech, ITC-18) at 100 μs resolution, and simultaneously the raw data signals were also acquired with a Plexon Recorder (filtered 3Hz to 8kHz, sampled at 40kHz) and streamed to hard disk for later analysis. Single-units were isolated using a window discriminator (Frederick Haer) and displayed on a delay-triggered digital oscilloscope. When recording with multielectrodes, spikes from one selected channel were analyzed online and used to guide the recording protocol (below). A photocell (TAOS, TSL12S) was used for temporal registration of stimulus onset/offset timing and spike recordings.

A manually controlled bar-shaped stimulus was used to search for neural signals, and to determine location of the receptive field, ocular dominance, eccentricity and approximate optimal orientation. The CRT monitor was centered on the neuron’s receptive field, and the non-dominant eye occluded. Drifting sinewave luminance gratings were used to measure the neuron’s luminance passband (spatial frequency, orientation and temporal frequency tuning), with each stimulus condition randomly interleaved and repeatedly presented for 10-20 times. Then the neuron’s optimal LM grating was presented in small circular patches in different locations on the
screen to more accurately map the receptive field, and the screen re-centered if necessary. To
measure the size of the receptive field and check for surround suppression, the optimal LM
grating was presented in circular patches of varying sizes centered on the receptive field.

As an initial assessment of responsiveness to second-order stimuli, responses to drifting
CM gratings were recorded with a stationary carrier, an envelope orientation at the neuron’s
optimal luminance orientation, an envelope spatial frequency equal to or lower than the neuron’s
optimal luminance spatial frequency and an envelope temporal frequency slightly lower than the
neuron’s optimal luminance temporal frequency (Zhou & Baker, 1996; Mareschal & Baker
1999). A series of carrier spatial frequencies were tested, ranging from values near the screen
resolution to the neuron’s luminance passband, to find an optimal carrier spatial frequency. We
classified a neuron as second-order responsive if it gave significant responses compared to
spontaneous activity (t-test) at relatively high carrier-spatial frequencies that were well outside
the luminance passband of the neuron, and if this spatial frequency tuning was band-pass. This
condition of band-pass tuning ensured that the neuron’s response was genuinely second-order
and not due to a nonlinearity in the screen which might give rise to a luminance signal at the
envelope spatial frequency (Zhou and Baker 1994). If the neuron was classified as second-order
responsive, then responses to velocity modulation (VM) gratings were recorded by testing a
series of carrier spatial frequencies, with envelope orientation fixed to the neuron’s optimal
luminance orientation. If a neuron responded significantly to VM gratings, then envelope
orientation tuning was measured using the neuron’s optimal carrier spatial frequency, with
carrier orientation always kept perpendicular to the envelope orientation. The temporal frequency
of the drifting carriers in VM gratings was then varied, to study carrier temporal tuning
properties. The temporal frequency response for the carrier of CM gratings was also obtained for comparison with the VM results.

**Analysis**

Neurons were classified as either simple or complex type by measuring the ratio of modulated (first harmonic) to mean responses \((F1/(F0 - \text{spont})\), or “AC/DC ratio”) to the neuron’s optimal LM grating. If the ratio was greater than one, the neuron was classified as a simple type cell, otherwise it was classified as a complex cell (Skottun et al. 1991). Neuronal responses used in the formulae below had spontaneous activity subtracted from them.

Spatial frequency tuning curves were fit with a gaussian function (DeAngelis et al. 1994) to obtain an estimated optimal spatial frequency,

\[
R(sf) = ke^{-\frac{(sf-SF_{opt}/\alpha)^2}{2}} + Ro
\]  

where \(k\), \(SF_{opt}\), \(\alpha\), \(Ro\) are free parameters and \(R\) represents neuronal response at spatial frequency \(sf\). A bootstrap resampling method (Efron and Tibshirani 1993) was used to estimate 95% confidence intervals for the obtained optimal spatial frequency \((SF_{opt})\) value.

For orientation tuning curves, circular variance (CV) was calculated as an index of tuning bandwidth (Marida 1972).

\[
CV = 1 - \frac{\left| \sum \mathcal{R}_k \exp(i2\theta_k) \right|}{\sum \mathcal{R}_k}
\]  

where \(\mathcal{R}_k\) represents neuronal response at orientation \(\theta_k\). Circular variance ranges from zero (sharp tuning) to unity (isotropic tuning). Optimal orientation was estimated as:
\[ \text{Ori}_{\text{opt}} = \arg \left( \frac{\sum_k R_k \exp(i2\theta_k)}{\sum_k R_k} \right) \]  

where \( \arg \) denotes angular component of a complex number.

Motion direction selectivity of a neuron was measured by a direction selectivity index (DSI),

\[ DSI = \frac{(R_p - R_N)}{(R_p + R_N)} \times 100\% \]  

where \( R_p \) is response of the neuron to its preferred direction of motion and \( R_N \) is response to its non-preferred direction. DSI ranges from 0 % (non-directional) to 100 % (completely directional).

Neurons’ responses to a series of carrier temporal frequencies in both directions of motion were tested for CM and uni-directional VM gratings. The extent to which these data revealed direction selectivity to carrier motion was summarized by a symmetry index (SI),

\[ SI = 1 - \frac{\sum_k |R_k - R_{-k}|}{\sum_k |R_k + R_{-k}|} \]  

where \( R_k \) is response of the neuron to VM or CM gratings with carrier drifting at ‘k’ Hz and \( R_{-k} \) is response to stimuli with carrier drifting at ‘k’ Hz in the opposite direction. SI would be 0 if the neuron responds only to one direction of carrier motion and not to the other (direction selective), and it would be 1 if the neuron responds equally to both directions of carrier motion (non-direction selective).

Proportional decline in response of a neuron at high temporal frequency compared to its optimal response is given by a fall-off index (FI),

\[ FI = \left( \frac{R_{hi}}{R_{max}} \right) \]
The fall-off index was calculated for responses of a neuron to LM gratings as well as VM and CM gratings. In the case of LM gratings, $R_H$ is response to gratings drifting at 16 Hz and $R_{\text{max}}$ is response at the optimal temporal frequency. In the case of VM and CM gratings, $R_H$ is response to VM or CM gratings with carrier drifting at 16 Hz, and $R_{\text{max}}$ is response to the same grating at its optimal carrier temporal frequency. FI ranges from 0 (response falls to spontaneous at 16 Hz) to 1 (optimal response at 16 Hz over the measured range of 0-16 Hz).

To evaluate whether a neuron exhibited similar preference for two kinds of gratings, Pearson’s correlation coefficient was calculated for scatter plots comparing optimal carrier spatial frequencies for CM & VM gratings and optimal orientations for LM gratings and envelope of VM gratings. Also, nonparametric, paired comparisons (Wilcoxon signed rank test) were performed.

For recordings with multi-electrodes, spikes from only one channel were analyzed online to construct tuning curves. In later offline analysis, spikes from other channels were detected and classified using the Offline Sorter software (V2.8.8, Plexon). Spikes were sorted using the “semi-automatic K-means” algorithm, and only clearly separable clusters of spikes were classified as single units. Isolated neurons from these channels were included in further analysis only if they showed very similar tuning to orientation and spatial frequency of LM gratings compared to the neuron recorded online. In some cases for recordings with single channel electrodes, offline sorting of spikes was performed to correct misclassifications by the window discriminator, and to isolate and assess spikes with lower amplitude.

**Model:**

To explore to what extent the model scheme in Fig. 1E-F could provide an understanding of the general features of these neuronal responses, we constructed a computer simulation in Matlab.
The architecture of the model (Fig. 1E) consists of two parallel processing pathways, a linear filter F0 responding to luminance (first-order) stimuli and a nonlinear pathway (F1-R-F2, Fig. 1F) processing non-luminance (second-order) stimuli.

Filter F0 is a spatiotemporal filter constructed by taking a dot-product of each frame of the stimulus with a gabor spatial filter, to produce a temporal signal that is then convolved with a temporal filter (Adelson & Bergen, 1985) and finally this signal is half-wave rectified to give a simple cell-like modulated response.

\[
\hat{f}_{\text{lin}}(t) = \frac{(k \ast t)^n e^{-kt}}{n!} - \frac{(k \ast t)^n}{n + 2}
\]  

(7)

where \( n = 2 \), \( t \) represents time and \( k \) is a constant calculated by the following equation.

\[
k = t m_{\text{scl}}^2 \ast f_{\text{opt}} \ast 1000
\]  

(8)

Where \( t m_{\text{scl}} \) is a time scale factor defined as milliseconds/frame, and \( t f_{\text{opt}} \) is the optimal temporal frequency of the filter.

The first stage of the nonlinear pathway contains a pool of spatiotemporal filters F1, implemented by convolving each spatial frame stimulus with a gabor spatial filter, followed by temporal convolution with a monophasic temporal filter (Watson and Ahumada, 1985).

\[
\hat{f}_{\text{lin}}(t) = (k \ast t)^n e^{-kt}
\]  

(9)

where \( n = 0 \), \( t \) represents time and \( k \) is given by equation 8 where \( t f_{\text{opt}} \) is set to 1Hz.

Each of these temporal responses is then full-wave rectified, and summed by a spatiotemporal filter F2 which is constructed exactly the same as filter F0. The action of filter F2 on the rectified signals is also implemented as a dot product in the same manner as in the linear pathway and then output of filter F2 is half-wave rectified to give a simple cell-like modulated response.
Finally the temporal output signals of these two pathways are summed to give the final output of the model. Note that the output of this model is an analog signal representing average spike frequency as a function of time rather than discrete spiking events.

We did not try to fit parameters of these filters to the data from individual neurons, but instead employed a generic model with fixed values of parameters, because our aim here was to explore to what extent the nonlinear FRF model that has been proposed to process second-order stimuli (Baker and Mareschal, 2001) can provide some understanding of these data. Model responses were measured to the same stimuli used in the experimental recordings of the neuronal responses. The parameters of the model were as follows: for filter F0, spatial frequency of the gabor was 0.08 cpd, spatial bandwidth was 1.5 octaves, aspect ratio (defined as the ratio of the filter’s axial length to cross width) was 1, orientation was 0 degrees and optimal temporal frequency was 2 Hz. For filter F1 spatial frequency of the gabor was 1.6 cpd, spatial bandwidth was 1.5 octaves, aspect ratio was 1, orientation was 90 degrees and the parameters of the monophasic temporal filter were n = 0 and t_{f_{opt}} = 1 Hz, which were chosen to provide selectivity to carrier temporal frequency of CM gratings roughly like those shown by neurons (Fig. 8 C-H). The parameters for filter F2 were identical to those for filter F0.
Results

For this study, we recorded from 115 Area 18 neurons in 13 cats. Out of these, 64 (55 %) were classified as second order responsive neurons, as they responded significantly to contrast modulation (CM) gratings and showed band-pass tuning to its carrier spatial frequency. These second order responsive neurons were further tested with velocity modulation (VM) gratings (motion-defined) of two types, uni-directional and bi-directional. Carrier spatial frequency and envelope orientation tuning were also measured using uni-directional VM gratings, which gave stronger responses than bi-directional VM gratings.

Carrier spatial frequency selectivity

Previous studies have demonstrated that second-order responsive neurons in Area 18 show narrow band-pass tuning to carrier spatial frequency of CM gratings well outside their luminance passband (e.g. Zhou and Baker 1996; Mareschal and Baker 1998; Tanaka and Ohzawa 2006; Rosenberg and Issa 2011). Thus we wanted to see if these neurons would also show similar selectivity to carrier spatial frequency of VM gratings. The neuron in Fig.2 showed bandpass tuning to luminance gratings (0.02-0.2 cpd, with a peak response at 0.08 cpd). The CM gratings were tested with varying carrier spatial frequencies, with envelope orientation fixed at the neuron’s optimal luminance orientation. This neuron showed band-pass tuning to carrier spatial frequency with a response peak at approximately 0.8 cpd. At this high spatial frequency range, the carrier signals were well beyond the neuron’s luminance resolution, and hence it was classified as second-order responsive neuron. This neuron was further tested with VM gratings for varying carrier spatial frequencies, with envelope orientation fixed at the neuron’s optimal luminance orientation. This neuron showed similar band-pass tuning to carrier spatial frequency
as shown for CM gratings, again with a response peak at approximately 0.8 cpd. In this plot the
neuron’s responses are shown on different scales for carrier spatial frequency tuning and LM
spatial frequency tuning, since this neuron’s LM grating response was much stronger than those
to the CM and VM gratings.

The scatter plot in Fig. 3A shows a given neuron’s optimal carrier spatial frequency for
VM gratings against that for CM gratings for 30 neurons (9 simple and 21 complex). Points on
the scatter plot are highly correlated ($r = 0.7$, $p < 0.0001$, 95% CI [0.46-0.85]) and most of the
points on the scatter plot lie close to the equality line, indicating that carrier spatial frequency
tuning is similar for both types of gratings. Only 1 neuron out of 30 had significantly different
optimal carrier spatial frequencies for two stimuli (bootstrap method, C.I. = 95%). Fig. 3B shows
a histogram plotted for differences in optimal carrier spatial frequencies between CM and VM
gratings, which is centered on zero and mean of these differences is 0.15 cpd. This mean is small
compared to the mean spatial frequency tuning bandwidths for CM (BW = 0.47 cpd) and VM
(BW = 0.62 cpd).

**Envelope orientation selectivity**

To assess “form-cue invariance” for luminance- and motion-defined boundaries, like that
previously demonstrated for luminance- and contrast-defined boundaries (Mareschal and Baker
1998a) we measured orientation selectivity of neurons to the envelope of “uni-directional” VM
gratings, for comparison with that of LM gratings. Fig. 4A shows responses of a typical neuron
that was tuned to orientation as well as motion direction of luminance gratings ($Ori_{opt} = 97$ deg,
DSI = 0.94). This neuron also showed similar orientation tuning for the envelope of VM gratings
(Fig. 4B) ($Ori_{opt} = 91$ deg), and was also direction-selective though to a somewhat smaller
degree (DSI = 0.55). For this neuron we also measured envelope orientation tuning for CM
gratings (Fig. 4C) ($Ori_{opt} = 96.1$, DSI = 0.37 deg), which showed tuning very similar to that for
VM gratings.

The scatter plot in Fig. 5A shows a given neuron’s optimal envelope orientation for VM
gratings against optimal orientation for LM gratings for 26 neurons (10 simple and 16 complex).
Points on the scatter plot are highly correlated ($r = 0.95$, $p < 0.0001$, 95% CI [0.89-0.98]) and
most of the points lie close to the equality line. 58 % (15/26) of the neurons had optimal
orientation differences less than 15 degrees, with a maximum orientation difference of 37
degrees. There is no significant difference between a given neuron’s optimal orientation for LM
gratings and optimal envelope orientation for VM gratings (Wilcoxon signed rank test, $p = 0.34$).
The histogram in Fig. 5B shows that differences in optimal orientations are very small (mean = -
4.4 deg). The scatter plot in Fig. 5C shows a given neuron’s envelope orientation circular
variance (CV) for VM gratings against orientation circular variance (CV) for LM gratings. Most
of the points (23 out of 26) lie above the equality line, indicating broader tuning for VM
compared to LM gratings. The CV for VM gratings is significantly greater than for LM gratings
(Wilcoxon signed rank one-tailed test, $p < 0.0001$). The scatter plot in Fig. 5D shows a given
neuron’s direction selectivity index for motion direction of envelope of VM gratings vs. motion
direction of LM gratings. Most of the points (18 out of 26) lie in the first quadrant suggesting
that most neurons preferred the same direction of motion for both kinds of stimuli. DSI for LM
gratings (mean = 0.52) is significantly greater than for VM gratings (mean = 0.23) (Wilcoxon
signed rank one-tailed test, $p = 0.0008$), suggesting that neurons had weaker direction selectivity
for VM compared to LM gratings.
Note that if a neuron was responding to the carrier motion, then the optimal envelope orientation for VM gratings would be orthogonal to that of LM gratings (since the carrier was always orthogonal to the envelope in these experiments), and the histogram in Fig. 5B would peak around 90 deg instead of 0 deg. However this is not the case - instead, these results demonstrate that neurons’ VM grating responses are to the envelope and not to the local motion of the carrier, and that they occur in a form-cue invariant manner.

**Carrier temporal frequency tuning**

A previous study (Mareschal and Baker 1998b) demonstrated that most of the second-order responsive neurons in cat Area 18 showed bandpass tuning to the temporal frequency of drifting envelopes of CM gratings. Neurons were systematically selective for lower envelope temporal frequencies of CM gratings compared to LM gratings. Responses of these neurons usually fall off to spontaneous activity above an envelope temporal frequency of about 10 Hz, while for luminance gratings, responses fall off around 16 Hz. In order to see if neurons show similar tuning properties to drifting carriers, we measured responses for CM, and both uni-directional and bi-directional VM gratings, for varying carrier temporal frequency. Note that the envelopes of CM, and both uni-directional and bi-directional VM gratings were drifted at a fixed temporal frequency that was slightly lower than the neuron’s optimal temporal frequency for LM gratings.

For uni-directional VM gratings, one carrier was always held stationary and other drifted at varying temporal frequencies in both directions. For some neurons, response decreased with increasing temporal frequency (Fig.6 C); response of some neurons increased with temporal frequency (E); some neurons responded equally to all temporal frequencies (F); a few showed band-pass tuning (D), and some showed no such particular pattern (G, H). Interestingly, almost
all neurons showed symmetric tuning, i.e. a similar response pattern for both directions of carrier motion, indicated by symmetry index values close to 1 in Fig. 6. The scatterplot in Fig. 9A shows a given neuron’s fall-off index for uni-directional VM gratings against that for LM gratings. Most of the points (19 out of 24) lie above the equality line, suggesting that for uni-directional VM gratings neuronal responses fall off at high temporal frequencies is relatively less than for LM responses. The fall-off index values are significantly greater for VM gratings compared to LM gratings (Wilcoxon signed rank one-tailed test, p = 0.0018).

For bi-directional VM gratings, the carrier gratings drifted with equal and opposite velocities. Fig. 7 shows carrier temporal frequency tuning for six neurons. The responses for most neurons in our sample decreased with increasing temporal frequency (e.g. Fig. 7 C-G), except for one (Fig. 7 H), which showed band-pass tuning. This was the only neuron that showed band-pass tuning in our sample of 19 neurons - all others gave low pass responses.

In most previous single-unit studies using CM stimuli, the carrier grating was always held stationary while the envelope was drifting (Zhou and Baker 1993, 1994; Mareschal and Baker 1998a, 1999; Song and Baker 2007; Rosenberg and Issa 2010). In our present study, in order to measure dynamic properties of early stages of a filter-rectify-filter model, we measured temporal frequency tuning for drifting carriers of CM gratings similar to the recent study (Rosenberg and Issa 2011). Since we wanted to compare this CM carrier temporal frequency tuning with the carrier tuning for VM gratings, we maintained the carrier orientation perpendicular to the envelope orientation for CM gratings, and varied carrier temporal frequency from 0 Hz to 16 Hz (0 Hz corresponding to a stationary carrier). Fig. 8 shows carrier temporal frequency tuning for six neurons. Consistent with (Rosenberg and Issa 2011) our sample neurons showed very diverse tuning, but most of them responded optimally to a stationary carrier, and response decreased with
increase in temporal frequency. Since the carrier of the CM gratings was stationary while evaluating neurons’ second-order responsivity, we might have inadvertently excluded in our sample any neurons that preferentially respond to CM gratings with drifting carriers (Rosenberg and Issa, 2011). However some neurons also responded quite well to very high temporal frequencies (Fig. 8 F, H). Similar to the results from uni-directional VM gratings, neurons usually showed symmetric tuning to carrier temporal frequency, i.e. the response pattern was similar to both directions of carrier motion (indicated by symmetry index values close to 1 in Fig. 8). The scatter plot in Fig. 9 B shows a given neuron’s fall-off index for CM gratings against that for LM gratings. There is no systematic relationship between fall-off indices for these two gratings, and the fall-off index values are not significantly different for CM and LM gratings (Wilcoxon signed rank two-tailed test, p = 0.98).

**Responsiveness to uni-directional and bi-directional gratings**

To assess how well CM-responsive neurons also responded to both uni- and bi-directional VM gratings, we compared response strength for the two stimuli at their optimal carrier temporal frequencies. The scatter plot in Fig. 10 A shows response strength to bi-directional vs uni-directional VM gratings for 19 neurons (5 simple and 14 complex) from carrier temporal frequency data like those in Figs.6 & 7. All the data points lie on or below the equality line, indicating that neurons responded more strongly to uni-directional VM gratings - this difference was significant (Wilcoxon signed rank one-tailed test, p < 0.0001). Even though all (19/19) of the second-order responsive neurons responded significantly to uni-directional VM gratings, some (9/19) failed to respond significantly (one tailed t-test) to bi-directional gratings. The histogram (Fig. 10 B) showing ratio of response strength to uni- and bi-directional stimuli demonstrates that for all neurons the ratio was less than or equal to unity, and for 16/19 neurons
the ratio was less than 0.6. The scatter plot and histogram show results only for those second-order responsive neurons that were tested with both types of VM gratings.

**Envelope spatial frequency tuning**

Conceivably VM gratings might be detected in two different ways - the dynamic discontinuities between the moving carriers (an “edge-based” processing) or by relative speeds of the carriers (a “region-based” processing) - see Discussion. We measured envelope spatial frequency tuning curves to uni- and bi-directional VM gratings and compared them to the tuning for illusory contours (IC), which are thought to be detected by edge-based processing (Song and Baker 2007). Interestingly, neurons were tuned to higher envelope spatial frequencies of uni-directional VM gratings compared to bi-directional ones. Neurons showed similar envelope spatial frequency selectivity for ICs and bi-directional VM gratings (Figs. 11 B). This similarity suggests that bi-directional boundaries are detected in a manner like ICs, i.e. by dynamic phase-discontinuities of the carrier textures along the boundary. However, neurons were selective for higher envelope spatial frequencies of uni-directional VM gratings compared to ICs, approximately two times the peak envelope spatial frequency for ICs (Figs. 11 A). Therefore, uni-directional boundaries may be detected primarily by differences in the speeds of textures between envelope half-cycles, i.e. region-based processing, with some additional contribution from dynamic discontinuities, i.e. edge-based processing.

**Model simulation**

We simulated an FRF model that was proposed previously to explain responses of Area 18 neurons to CM gratings and ICs (Song & Baker, 2007), to see how well it might also account
for the responses of these neurons to VM gratings. The model parameters were not fit to a
particular neuron’s data, but rather a generalized model was constructed to simulate the typical
selectivity patterns of Area 18 neurons for CM gratings, which was then tested to see how well it
could predict selectivity patterns for VM gratings. Temporal parameters of the early filters were
selected to mimic typical neurons’ responses to carrier temporal frequency of CM gratings, such
that the model gave its best response to CM gratings with a stationary carrier and the response
decreased with carrier temporal frequency (Fig. 8B) - i.e. it gave a low-pass carrier temporal
frequency response. These filters were not selective for the direction of carrier motion, in
accordance with the neuronal responses (Fig. 8B; Rosenberg & Issa, 2011).

After fixing the parameters of the model based on CM responses, we tested its selectivity
to VM gratings and ICs. For uni-directional VM gratings, response of the model increased with
carrier temporal frequency (Fig. 6B). Though not all neurons’ responses were exactly like this
response pattern, most of the neurons’ responses did remain significantly above spontaneous
activity at high carrier temporal frequencies, as indicated by the fall-off indices in Fig. 9A. On
the other hand, the model’s response decreased with carrier temporal frequency for bi-directional
VM gratings (Fig. 7B), similar to the response pattern of nearly all the neurons in our sample.
And also similar to the neurons’ responses, the model responded less strongly to bi-directional
than to uni-directional VM gratings. Interestingly, the model showed similar envelope spatial
frequency selectivity for ICs and bi-directional VM gratings (Fig. 11B), while it was selective for
an envelope spatial frequency of uni-directional gratings that was twice that for ICs (Fig. 11A).
These VM and IC envelope spatial frequency selectivity results were similar to those for
neurons, thus supporting our hypothesis that bi-directional gratings are detected by “edge-based
processing” similar to ICs, while uni-directional gratings are detected predominantly by “region-based processing”.
Discussion

This study has demonstrated that neurons in early visual cortex can respond to motion-defined contours. These neurons were selective to similar orientations of luminance- and motion-defined contours, as well as similar directions of motion (form-cue invariance), although the strength of direction selectivity was weaker for motion-defined contours compared to luminance contours. These neurons were also selective for spatial frequency of the carrier gratings used for constructing motion-defined contours. This carrier selectivity was very similar to the selectivity shown for the carrier of contrast-defined contours. These findings suggest that both kinds of contours are detected by the same nonlinear neural mechanism. Responses to both contrast- and motion-defined boundaries often extended to quite high temporal frequencies of drifting carrier gratings to which most cortical neurons fail to respond when tested with luminance gratings. But for a given neuron, tuning was similar for both directions of carrier motion.

Sinusoidal grating carrier

It might seem counterintuitive to use sinusoidal gratings as a carrier instead of random dot texture patterns used in previous studies (Chaudhuri and Albright 1997; Mysore et al. 2006; Sary et al. 1993, 1995; Zeki et al. 2003), as these patterns look more similar to texture patterns present in the real world. However a sinusoidal grating carrier with spatial frequency outside a neuron’s luminance passband provides powerful advantages in ruling out simple linear/luminance artifacts. Firstly, it ensures that the responses are not mediated by the same linear filter thought to process luminance gratings. A random dot texture pattern however, is broadband in spatial frequency and some of its energy might fall within the luminance passband of a neuron, giving rise to a linear response. Secondly, these neurons show narrow carrier spatial
frequency tuning, and this result rules out the possibility that their responses are mediated by early nonlinearities of the retina or CRT because such nonlinearities would not predict selectivity for carrier spatial frequencies.

However VM stimuli with sinusoidal grating carriers introduce an ambiguity as to what the neuron is actually responding to. If a snapshot image of VM stimuli is taken at some moment in time (Fig. 1 C) then it looks like an illusory contour with phase-discontinuities between two carriers. But measuring envelope spatial frequency tuning to VM gratings as described below, we were able to disambiguate between responses to phase discontinuities and difference in speeds of carriers. Note that in natural scenes, shear motion between textures will give rise to both of these cues: phase discontinuities and relative speed.

**Neural Mechanism**

A model consisting of two parallel signal-processing streams (Fig. 1.E) has been proposed to explain responses of cortical neurons to first- and second-order stimuli (Zhou and Baker 1993; Mareschal and Baker 1999; Song and Baker 2007). The first stream consists of a conventional coarse spatial scale linear filter (F0) selective for orientation, direction of motion and spatial frequency of luminance gratings. The second stream consists of a nonlinear Filter-Rectify-Filter (FRF) model (Fig. 1.F) that can explain responses of neurons to second-order stimuli such as contrast- and texture-defined boundaries (Chubb and Sperling 1988; Wilson 1999; Landy & Graham, 2004). This FRF model is composed of two linear filtering stages that are connected by a nonlinearity (eg. rectification). The first stage consists of small-scale spatial filters (F1) that are selective for high spatial frequencies of the carrier. The outputs of these early filters are rectified and pooled by a coarse spatial scale late filter (F2), which is selective for envelope orientation, direction of motion and spatial frequency. Filters F0 and F2 have similar
preference for orientation and direction of motion, but spatial frequency selectivity is coarser for F2 compared to F0. Here we explore whether the FRF model proposed to explain responses to contrast- and texture-defined boundaries could also explain responses of Area 18 neurons to motion-defined boundaries.

If motion-defined boundaries are processed by a common FRF-like mechanism, then neurons should show similar tuning properties for carrier and envelope of motion- and contrast-defined boundaries. In processing of motion-defined boundaries by an FRF model, filters F1 could act as local motion detectors whose outputs will be rectified and pooled by filter F2. If the same F1 filters are used for processing motion- and contrast-defined boundaries, then neurons should have similar carrier spatial frequency selectivity for both kinds of boundaries. Our results demonstrate that a given neuron is indeed selective for similar carrier spatial frequency for motion- and contrast-defined gratings (Fig. 2).

Such an FRF model could detect VM grating stimuli in two ways. The neuron could be responding to differences in the speeds of drifting carriers between adjacent half cycles of VM gratings (a region-based processing) and/or to the dynamic discontinuities (carriers in adjacent half-cycles moving in and out of phase with one another) along the boundaries (an edge-based processing) similar to illusory contours (Song and Baker 2007). Since dynamic phase discontinuities are present in both uni-directional and bi-directional VM gratings, an FRF model could produce an edge-based response to both stimuli. For this model to respond in a region-based manner to uni-directional VM gratings, its early F1 filters must respond differently to stationary and moving carriers. For bi-directional VM gratings, an FRF model will give a region-based response only if its early filters can distinguish between carriers drifting with equal speeds in opposite directions, which is possible only if the early filters are selective for motion direction.
We measured temporal tuning properties of early filters by systematically varying the temporal frequency of a drifting carrier grating for contrast-defined boundary stimuli (Fig. 8). The neurons’ responses were symmetric for both directions of carrier motion, suggesting that early filters of the FRF model are not direction selective. For most neurons, response peaked when the carrier was held stationary, and gradually declined with increasing carrier temporal frequency. Since early filters of the FRF model are not direction selective, the FRF model would predict (Fig. 6B) symmetric carrier temporal frequency tuning for both directions of carrier motion for uni-directional VM gratings, which our results (Fig. 6 C-H) largely demonstrate. The model would also predict that responses to uni-directional VM gratings would be mediated by both edge-based processing (because of dynamic discontinuities) and region-based processing (because early filters give different response to stationary and moving carrier). Responses to bi-directional VM gratings would be mediated by edge-based processing only, and not by region-based processing (because early filters are not direction selective) as shown in (Fig. 12).

To test these predictions about mechanism, we measured envelope spatial frequency tuning for both uni-directional and bi-directional VM gratings, and compared it with envelope spatial frequency tuning for illusory contours (ICs). ICs are thought to be detected in an edge-based manner (Wilson 1999; Song and Baker 2007), so if a VM grating is also detected in an edge-based manner it will show the same tuning for envelope spatial frequency as an IC. But if a VM grating is detected by region-based processing (like contrast-defined boundaries (Song and Baker 2007)) it will be tuned to an envelope spatial frequency twice that of ICs, as there are two phase-discontinuity edges in one envelope cycle of an IC. Our results (Fig. 11) from both neurons and model simulations show that bi-directional VM grating responses are tuned to the same envelope spatial frequencies as illusory contours. But uni-directional VM grating responses
are tuned to spatial frequencies approximately twice the optimal for illusory contours. This result suggests that uni-directional gratings are detected by a mixture of both region-based and edge-based processing. Also our results (Fig. 10) show that responses of neurons are stronger to uni-directional than to bi-directional gratings - this could be because uni-directional gratings are simultaneously detected by both region-based and edge-based processing, while bi-directional gratings are detected by edge-based processing only.

Results from our model simulations also predict that neurons’ fall-off indices for carrier temporal frequency to uni-directional and bi-directional gratings would be different (Fig. 7B, 8B), which is demonstrated by neurons responses in (Fig. 7, 8). But it should be noted that this model is not accurate in explaining all the details of individual neuron’s responses, particularly the diversity of temporal responses (Fig. 6, 7, 8) – to do so would require a more elaborate model architecture and additional model parameters, whose estimation would be beyond the aims and scope of this study. The purpose of this model is just to demonstrate in a general way how a simple FRF model could produce patterns of selectivity to different second-order stimuli that are similar to the neuronal data, particularly different carrier temporal frequency fall-off indices to different stimuli and differing envelope spatial frequency selectivities.

The neural substrates for the elements of the filter-rectify-filter model are not known with certainty. Earlier it was proposed that Area 17 neurons could be the basis for the early filters (F1) due to their selectivity for high spatial frequency, and the orientation selectivity shown for the carrier of CM gratings which seemed to rule out a subcortical substrate (Mareschal and Baker, 1998). However a recent study (Rosenberg et al., 2010) showed that LGN Y cells also respond selectively to CM gratings - most interestingly, the nonlinear subunits of Y cells exhibited carrier orientation tuning similar to that seen in Area 18 neurons. These results suggest that nonlinear
subunits of Y cells could provide the early filters (F1), with subsequent cortical summation of Y
cell afferents providing envelope orientation and spatial frequency tuning (F2) observed in Area
18 neurons. Also consistent with this idea are the very good responses to quite high carrier
temporal frequencies in both LGN Y cells (Rosenberg et al, 2011) and Area 18 neurons (Fig. 8;
Rosenberg and Issa, 2012), whereas Area 17 neurons fail to respond at high temporal frequencies
(Movshon et al, 1978). On the other hand, carrier spatial frequency tuning bandwidth is very
broad in LGN Y cells compared to that in Area 18 neurons (Rosenberg et al, 2010), so it is not
yet clear to what extent the early filtering (F1) might be accounted for at the geniculate level.

In summary, our results are consistent with the idea that these VM grating responses can
be understood in terms of the same FRF model previously proposed for responses to CM and IC
stimuli, suggesting a common neural mechanism for all these second-order stimuli.

**Cue-invariance**

For the visual system to perform figure-ground segregation, it should be able to delineate
an object from a background, which might be the same with respect to all cues except one, e.g.
color. However that object should also stand out if it appears against a different background with
the same color but a different texture. In order to perform this task, information about presence
of a particular cue is not important, but rather the contrast between cues that distinguish an object
from its background is of primary importance, regardless of the nature of the cues. So the visual
system should be able combine information across different cues to perform figure-ground
segregation, i.e. the segmentation mechanism should be “form-cue invariant” (Albright 1992).
This strategy of form cue-invariance is computationally economical, and can help resolve
perceptual ambiguities when multiple cues are present. In addition it might be important for
shape recognition and shape constancy.
Neurons in the early visual cortex have been previously shown to respond in a form-cue invariant manner to stimulus attributes such as orientation and motion direction of boundaries. Neurons in cat Area 17 (Zhou and Baker 1993, 1996), cat Area 18 (Zhou and Baker 1993, 1996; Mareschal and Baker 1998a, 1998b; Leventhal et al. 1998; Zhou et al. 2001; Song and Baker, 2006, 2007; Tanaka and Ohzawa, 2006), primate V1 (Chaudhuri and Albright, 1997) and primate V2 (e.g., von der Heydt et al. 1984; von der Heydt and Peterhans 1989; Leventhal et al. 1998; Lui et al. 2005) have been shown to respond in a form-cue invariant manner to the orientation of luminance and non-luminance boundaries.

A long-standing concern has been the possibility that second-order responses might be mediated by simple early nonlinearities in the display device or the photoreceptors, or by luminance signals in the stimuli, without any implication of a specialized mechanism (e.g., filter-rectify-filter) to explain the form-cue invariance. However the demonstration of carrier-tuned responses to stimuli with sinewave grating carriers in earlier studies (e.g. Zhou & Baker, 1994; Mareschal & Baker, 1998) as well as here (Fig. 2) make such explanations seem very unlikely.

One group has failed to find carrier-tuned second-order responsive neurons in cat Area 18 and primate V2 (El-Shamayleh and Movshon, 2011), but other laboratories have consistently replicated such tuned responses in cat Area 18 (Ohzawa et al; Issa NP et al) and recently also in primate V2 (Li et al, 2011).

In this study we demonstrate that neurons previously shown to selectively respond to luminance-, texture- and contrast-defined boundaries can also display form-cue invariant responses to motion-defined boundaries. These neurons showed similar orientation selectivity for luminance- and motion-defined boundaries, but direction selectivity was weaker to motion-defined boundaries, similar to what was found previously for contrast- and texture-defined
boundaries (Song and Baker 2006, 2007). This form-cue invariant orientation tuning in early visual cortex could be utilized by higher brain areas like MT, V4 and IT that have been reported to respond to motion-defined patterns in a form-cue invariant manner (Albright 1992; Mysore et al. 2006; Sary et al. 1993).

This study suggests that specific processing of motion-defined boundaries begins in early visual areas. Further, it suggests that these boundaries are processed by a common nonlinear mechanism that also mediates response to other second-order stimuli such as contrast-and texture-defined boundaries, and it could be the basis for form cue-invariance to these stimuli.
Acknowledgements

We would like to thank Guangxing Li for providing the software for curve-fitting and Plexon data file analysis. Also, we would like to thank Guangxing Li, Vargha Talebi and Lynda Domazet for assistance with the experiments.

Grants

This work was supported by the Canadian Institutes of Health Research grants MA-9685 and MOP-119498 to C.L. Baker.

Disclosures

No conflicts of interest, financial or otherwise, are declared by the authors.
References


Rogers B and Graham M. Motion parallax as an independent cue for depth perception. 


Rosenberg A, Husson TR and Issa NP. Subcortical representation of non-Fourier image features. 


Rosenberg A and Issa NP. The Y cell visual pathway implements a demodulating nonlinearity. 


Scott-Samuel NE and Georgson MA. Does early non-linearity account for second-order motion? 


Figure Legends

Figure 1. Four types of grating stimuli used in this study, and a model. (A) Luminance modulation (LM) sinusoidal grating with vertical orientation. (B) Contrast modulation (CM) grating with vertically oriented sinusoidal envelope that modulates the contrast of a horizontal high spatial frequency carrier grating. (C) “Uni-directional” velocity modulation (VM) grating with vertically oriented squarewave envelope that modulates the velocity of a horizontal high spatial frequency carrier grating. For “uni-directional” VM, the carrier in half of the envelope cycles is stationary, and in the other half it drifts with a specified temporal frequency. (D) “Bi-directional” velocity modulation (VM) grating is constructed similarly to “uni-directional” VM, except that the carrier in alternate half cycles of the envelope drifts with equal speeds but in opposite directions. In the stimulus images here and in the following figures, carrier motion is indicated by thin, white arrows, while envelope motion is depicted by thick, gray arrows. (E) Schematic model for neuronal responses, in which first- and second-order responses are mediated by separate, parallel pathways. The top pathway is a coarse spatial scale linear filter (F0) that would be responsive to conventional LM gratings. The bottom pathway mediates nonlinear processing of CM and VM gratings. (F) Non-linear FRF model that responds selectively to CM and VM gratings. The first stage of the model consists of small scale filters (F1) that are selective for high spatial frequency carriers. The outputs of these F1 filters are rectified and pooled by a late coarse scale filter, F2, which would be selective for the envelope of CM and VM gratings and would have similar spatio-temporal properties as the F0 filter.

Figure 2. Responses of a typical neuron to luminance modulation (LM), contrast modulation (CM) and velocity modulation (VM) gratings. Neuronal responses to LM gratings are shown as a
function of spatial frequency measured at grating’s optimal orientation. Snapshots of luminance
gratings with two different spatial frequencies are shown at the top. Optimal luminance spatial
frequency for this neuron was 0.08 cpd, and neuronal response fell to spontaneous activity
(dashed line) at 0.3 cpd. Responses of the same neuron are measured to CM and VM gratings as
a function of carrier spatial frequency. Carrier spatial frequency tuning for both gratings was
very similar, with peaks around 0.8 cpd, much greater than the optimal luminance grating spatial
frequency of 0.08 cpd. Snapshots of CM and VM gratings with two different carrier spatial
frequencies are shown at the bottom.

**Figure 3.** Relationship between optimal carrier spatial frequency for VM and CM stimuli. (A)
Scatter plot showing neurons’ optimal carrier spatial frequency for VM gratings vs. CM gratings
for 30 neurons (21 complex and 9 simple cells). A given neuron’s optimal carrier spatial
garment for VM gratings is highly correlated with that for CM gratings ($r = 0.7$, $p < 0.0001$).
(B) Histogram showing ratios of optimal carrier spatial frequencies for CM and VM gratings in
octaves.

**Figure 4.** Orientation tuning of a typical neuron to LM gratings and envelopes of VM and CM
gratings. In these polar plots distance from the origin indicates neural response (spikes/sec);
angular subtense represents envelope orientation (0-360 degrees). Snapshots of LM, VM and CM
gratings at three different orientations are shown next to the polar plots. Optimal orientation,
circular variance (CV) and direction selectivity index (DSI) are shown at the bottom of each
polar plot. This neuron showed similar orientation tuning and direction selectivity for all three
stimuli, i.e. form-cue invariance.
Figure 5. Relationship between orientation and direction selectivity for VM and LM stimuli, for a sample population of neurons. (A) Scatter plot showing neurons’ optimal envelope orientation for VM gratings vs. optimal orientation for LM gratings for 26 neurons (10 simple and 16 complex cells). A given neuron’s optimal envelope orientation for VM gratings is highly correlated with its optimal orientation for LM gratings ($r = 0.95$, $p < 0.0001$). (B) Histogram showing differences between optimal orientations for VM and LM gratings. (C) Scatter plot showing a given neuron’s envelope orientation circular variance (CV) for VM gratings vs. orientation circular variance (CV) for LM gratings. Most of the points (23 out of 26) lie above the equality line, indicating broader tuning for VM compared to LM gratings. (D) Scatter plot showing a given neuron’s direction selectivity index for motion direction of envelope of VM gratings vs. motion direction of LM gratings. Most neurons (18 out of 26) preferred the same direction of motion for LM and VM gratings. The remaining neurons that preferred opposite directions were weakly direction selective to at least one of the two gratings.

Figure 6. Carrier temporal frequency responses to "uni-directional" VM gratings (A), for the model (B) and six neurons (C-H). Different neurons showed diverse tuning properties to carrier temporal frequency, but most responded significantly to a broad range of temporal frequencies tested, with similar tuning for both directions of carrier motion. Responses of the model (whose parameters were chosen based on CM responses - Fig 8) increased with carrier temporal frequency. Negative values of carrier temporal frequency signify carrier motion in opposite direction. Dashed lines indicate spontaneous activity. Symmetry index (SI) values represent symmetry of responses to both directions of carrier motion, and fall-off index (FI) values
represent relative fall in response at carrier temporal frequency of 16 Hz compared to maximum response.

**Figure 7.** Carrier temporal frequency responses to "bi-directional" VM gratings (A), for the model (B) and six neurons (C-H). Responses of the model (B) decreased with increasing carrier temporal frequency. Similar to the model responses of the neurons typically decreased with increasing carrier temporal frequency (C-G), while neuron (H) showed band-pass tuning. Dashed lines indicate spontaneous activity. Fall-off index (FI) values are shown at the top of each plot.

**Figure 8.** Carrier temporal frequency responses to CM gratings (A), for the model (B) and 6 neurons (C-H). Orientation of the carrier grating was kept perpendicular to the envelope orientation, and temporal frequency was varied from 0Hz (stationary carrier) to 16Hz in both directions. The model responded optimally to a stationary carrier and response decreased with increasing carrier temporal frequency. Neurons responded similarly to the model (C-F), though some neurons responded equally well to drifting carriers (G, H). Neurons showed similar tuning for both directions of carrier motion. Negative values of carrier temporal frequency indicate motion in opposite direction. Dashed lines indicate spontaneous activity. Symmetry index (SI) and fall-off index (FI) values are shown at the top of each plot.

**Figure 9.** Comparison of fall-off index for "uni-directional" VM and CM gratings with LM gratings. (A) Scatter plot showing neurons’ fall-off index for “uni-directional” VM gratings vs. LM gratings for 24 neurons. Most points lie above the equality line (19/24), indicating that for VM gratings, neuronal responses fall off relatively less than LM responses at high temporal
frequencies. (B) Scatter plot showing neurons’ fall-off index for “bi-directional” VM gratings vs.
LM gratings for 22 neurons. There was no systematic relationship between fall-off indices for
these two gratings.

**Figure 10.** Comparison of peak response amplitudes to "uni-directional" and "bi-directional"
VM gratings. (A) Scatter plot showing a neurons’ maximum response (spikes/sec) to “bi-
directional” vs. “uni-directional” VM gratings for 19 neurons. Neurons responded more strongly
to “uni-directional” compared to “bi-directional” gratings. (B) Histogram showing ratio of
response strength to “bi-directional” gratings and “uni-directional” gratings. Ratio of response
strength was less than 0.6 for 16 out of 19 neurons (84.2%).

**Figure 11.** Comparison of neurons’ and model’s envelope spatial frequency (SF) tuning to
illusory contours (IC) with that to VM gratings. (A) Comparison of responses to IC and “uni-
directional” gratings. Responses from four neurons (lower graphs) show the preferred envelope
SF for “uni-directional” VM gratings was two times higher than that for ICs, as predicted by the
model (upper right graph). (B) Comparison of responses to IC and “bi-directional” gratings.
Responses from two neurons (lower graphs) show neurons’ preferred envelope SFs for “bi-
directional” VM gratings and ICs were the same, as also predicted by the model (upper right
panel).

**Figure 12.**
A schematic FRF model and its action on illusory contours (IC), and “bi-directional” and “uni-
directional” VM gratings. The late filter F2 is shown superimposed on full-wave rectified
responses of the early filters F1 for snapshot images of each stimulus, to show selectivity for
envelope spatial frequency and orientation. Note the “edge-based processing” for ICs and “bi-
directional” gratings, and “region-based” processing for “uni-directional” gratings, and the
corresponding difference in envelope spatial frequency selectivity.
Fig. 2
Fig. 5.

A

![Graph A](image)

B

![Histogram B](image)

C

![Graph C](image)

D

![Graph D](image)
Fig. 6

A

B

C

D

E

F

G

H

spikes/sec

VM_{uni} carrier TF (Hz)

spikes/sec

VM_{uni} carrier TF (Hz)

Si = 1; F1 = 1

Si = 0.95; Fl = 0.57

Si = 0.91; Fl = 0.44

Si = 0.92; Fl = 0.93

Si = 0.94; Fl = 0.81

Si = 0.86; Fl = 0.7

Si = 0.86; Fl = 0.29
Fig. 8

A

B

C

D

E

F

G

H

CM carrier TF (Hz)

CM carrier TF (Hz)

SI = 0.93; FI = 0.27

SI = 1; FI = 0.03

SI = 0.96; FI = 0.1

SI = 0.94; FI = 0.123

SI = 0.89; FI = 0.1

SI = 1; FI = 0.03

SI = 0.91; FI = 0.78

SI = 0.96; FI = 0.51

SI = 0.96; FI = 0.51

SI = 0.76; FI = 0

SI = 0.96; FI = 0.123

SI = 0.76; FI = 0

SI = 0.93; FI = 0.27

SI = 0.89; FI = 0.1

SI = 0.94; FI = 0.123

SI = 0.91; FI = 0.78

SI = 0.96; FI = 0.51

SI = 0.76; FI = 0

SI = 0.93; FI = 0.27

SI = 0.89; FI = 0.1

SI = 0.94; FI = 0.123

SI = 0.91; FI = 0.78

SI = 0.96; FI = 0.51

SI = 0.76; FI = 0

SI = 0.93; FI = 0.27

SI = 0.89; FI = 0.1

SI = 0.94; FI = 0.123

SI = 0.91; FI = 0.78

SI = 0.96; FI = 0.51
Fig. 9

(A) Graph showing the relationship between VM fall-off index and LM fall-off index for complex cells and simple cells.

(B) Graph showing the relationship between CM fall-off index and LM fall-off index for complex cells and simple cells.
Fig. 12

Early filter (F1) → Rectification → Later filter (F2)

IC  VM-Bi  VM- Uni