Behavioral Receptive Field for Ocular Following in Humans: Dynamics of Spatial Summation and Center-Surround Interactions

Frédéric V. Barthélemy, Ivo Vanzetta, and Guillaume S. Masson

Institut de Neurosciences Cognitives de la Méditerranée, Centre National de la Recherche Scientifique Unité Mixte de Recherche 6193, Aix-Marseille Université, Marseille, France

Submitted 1 February 2006; accepted in final form 7 March 2006

Barthélemy, Frédéric V., Ivo Vanzetta, and Guillaume S. Masson. Behavioral receptive field for ocular following in humans: dynamics of spatial summation and center-surround interactions. J Neurophysiol 95: 3712–3726, 2006. First published March 22, 2006; doi:10.1152/jn.00112.2006. Visual neurons integrate information over a finite part of the visual field with high selectivity. This classical receptive field is modulated by peripheral inputs that play a role in both neuronal response normalization and contextual modulations. However, the consequences of these properties for visuomotor transformations are yet incompletely understood. To explore those, we recorded short-latency ocular following responses in humans to large center-only and center-surround stimuli. We found that eye movements are triggered by a mechanism that integrates motion over a restricted portion of the visual field, the size of which depends on stimulus contrast and increases as a function of time after response onset. We also found evidence for a strong nonisodirectional center-surround organization, responsible for normalizing the central, driving input so that motor responses are set to their most linear contrast dynamics. Such response normalization is delayed about 20 ms relative to tracking onset, gradually builds up over time, and is partly tuned for surround orientation/direction. These results outline the spatiotemporal organization of a behavioral receptive field, which might reflect a linear integration among subpopulations of cortical visual motion detectors.

INTRODUCTION

The “receptive field” is a central notion in visual neuroscience: at the single-cell level, neurons integrate and respond to information over a restricted part of the visual field, with both high selectivity and sensitivity (Barlow 1953; Hartline 1938; Kuffler 1953). The receptive field’s functional properties can be characterized from the neuronal responses by a set of basic operators such as spatial summation and contrast-response functions (for a review see Albrecht et al. 2003; Carandini 2004). Furthermore, this (excitatory) receptive field is often modulated by visual inputs from its periphery (Allman et al. 1985). Such center-surround organization is involved in many fundamental tasks of visual processing such as input normalization (e.g., Bonin et al. 2005; Carandini et al. 1996; Heeger 1992; Schwartz and Simoncelli 2001), object motion segmentation (Allman et al. 1985; Nakayama and Loomis 1974), or contour detection (Field et al. 1993).

Several attempts have been made to measure the functional properties of neuronal receptive field properties, such as for motion perception (e.g., Anderson and Burr 1991; Chubb et al. 1989; Tadin et al. 2003; Watson and Turano 1995; Xing and Heeger 2001) and, to a lesser extent, for visuomotor transformations such as those underlying ocular tracking behavior (Heinen and Watamaniuk 1998; Miles et al. 1986). Ocular following responses are reflexive eye movements that, in both humans and monkeys, are initiated at ultrashort latencies (about 85 and 55 ms, respectively) and exhibit many of the properties attributed to the earliest stages of motion detection and integration (for reviews, see Masson 2004; Miles 1998). We and others have shown in humans that the earliest phase of ocular following is driven by luminance-defined motion (Masson et al. 2002) with a very high contrast gain (Masson and Castet 2002; Sheliga et al. 2005). Moreover, in monkeys, ocular following is initiated in response to motion presented in the central (<30°) visual field but can be positively modulated by an antagonistic peripheral motion (Miles et al. 1986), suggesting that the large-scale center-surround organization of visual motion mechanisms might have a functional role for tracking eye movements (Miles 1998). In humans, however, this center-surround modulation has so far been difficult to demonstrate in similar experiments (Gellman et al. 1990; Heinen and Watamaniuk 1998). Nevertheless, both ocular following (Gellman et al. 1990) and slow phases of optokinetic nystagmus (Abadi et al. 2005; Howard and Ohmi 1984) appear to be more efficiently driven by motion in the center of the visual field.

Here we introduce the concept of a large-scale “behavioral receptive field” that describes how a population of direction-selective neurons is organized to drive ocular following responses. First, we proceed to define such a functional entity with the help of three fundamental operators commonly used to characterize a receptive field (Carandini 2004): 1) a spatial summation function, 2) its dependency on contrast for central stimuli, and 3) its modulation by peripheral stimuli. Next, we explore those operators’ temporal dynamics and, finally, we show that these behavioral receptive field’s properties mimic many of the properties of neuronal receptive fields found at the earliest stages of the cortical visual motion pathway. Although further research is clearly needed to bridge the gap between the single-neuron level and the behavioral output, the functional characterization of such a “behavioral receptive field” provides theoretical constraints on the various suggested models and mechanisms of how local motion signals are integrated from neuronal populations to extract an efficient motor drive.
**METHODS**

**Subjects**

Five subjects (two naïve and the three authors, mean age: 34 ± 6 yr) participated in the study. All subjects were free of neurological diseases and underwent eye examination before participating in the experiments. All subjects had normal or corrected-to-normal acuity. All experiments were in accordance with the principles of the Declaration of Helsinki and followed the CNRS guidelines for conducting experiments in humans.

**Eye movements recording and visual stimuli**

The techniques used were described previously (Masson and Castet 2002; Masson et al. 2000). Eye movements were recorded using the electromagnetic search-coil technique (Robinson 1963) using coils embedded in a Silastin scleral ring placed in one eye after application of one to two drops of oxybuprocaine (Collewijn et al. 1985). Data acquisition, on-line control of the behavior, and stimulus triggering were controlled by a PC using the REX software package with the real-time QNX operating system (Hays et al. 1982). Voltage signals separately encoding horizontal and vertical positions of the right eye were low-pass filtered (Bessel, six poles, DC, 180 Hz) and sampled at 1 kHz, with a resolution of 16 bits.

Subjects were seated in a fiberglass chair, with chin and headrests, and faced a large (70 × 70°) vertical screen at a viewing distance of 1 meter. Visual stimuli were back-projected using a high-resolution Barco 809s video-projector (resolution: 1,280 × 1,024 at 76 Hz, frame duration: 13.2 ms). Visual stimuli were precomputed movies, generated using the HIPPS libraries (Landy et al. 1984) and stored into the memory of a SGI Fuel workstation. Visual workstation and experimental PC communicated throughout a serial port. Synchronization between the two computers was fully described elsewhere (Masson et al. 2000).

Center stimuli were always vertical sinusoidal luminance gratings drifting either rightward or leftward (spatial frequency: 0.27 cpi; temporal frequency: 10 Hz; speed: 37°/s). The motion directions of the central, driving, pattern were fully randomized across trials, temporal frequency: 10 Hz; speed: 37°/s). The motion directions of the central, driving, pattern were fully randomized across trials, temporal frequency: 10 Hz; speed: 37°/s). The motion directions of the central, driving, pattern were fully randomized across trials, temporal frequency: 10 Hz; speed: 37°/s).

Experiments were organized into several daily recording sessions of nearly 1 h duration, to collect usually around 150–200 trials for each condition. The signal-to-noise ratio (SNR) necessary to adequately resolve the responses was then achieved through averaging. The data from all sessions were pooled and analyzed off-line as described previously (Masson et al. 2000). Because ocular following was triggered in the close temporal vicinity of a saccade, all data shown here have the saccade-only condition subtracted to eliminate effects arising from post-saccadic drift.

Quantitative analysis was performed by measuring, for each trial, the changes in vertical (Δe|c) and horizontal (Δe|s) eye position over several 20-ms time windows. To achieve a complete description of the behavioral tuning for a given parameter (such as total contrast, stimulus diameter), we fitted the data with well-established descriptiv functions using the Marquardt–Lindberg algorithm (Matlab, The MathWorks, Natick, MA). The goodness of fits was estimated by computing a normalized Ch² value (Cavanaugh et al. 2002). First, to estimate the contrast at half-maximum response (Csign) and the slope (n) of the contrast-response functions of grating-driven tracking responses, we fitted the contrast-response function of the ocular following responses with a Naka–Rushton function (Albrecht and Hamilton 1982)

\[
R = R_{max} \left( \frac{c}{c + C_{sign}} \right)
\]

Here and in all subsequent formulae, C is contrast and R is the response (i.e., the change in position over a given 20-ms time window) to be fitted to the experimental data. Second, we assumed visual information to be integrated over the central, excitatory part of the visual field. This spatial summation function was expressed as an integral of a Gaussian error function (erf) (DeAngelis and Uka 2003; Pack et al. 2005)

\[
R(d) = A_e \text{erf} \left( \frac{d}{\omega} \right) + R_0
\]

and fitted to the responses R(d); d is the stimulus size, ω is the size of the excitatory part of the visual field (i.e., the width of the Gaussian function), R0 is the response amplitude offset, and A_e is the modulation amplitude. In some cases, largest stimulus radii resulted in a slight suppression of the amplitude of responses. To include such suppression, we also fitted these responses to a function of the integral of the difference between two Gaussians, with different amplitudes and widths for the excitatory and suppressive component

\[
R(d) = A_e \text{erf} \left( \frac{d}{\omega_e} \right) - A_s \text{erf} \left( \frac{d}{\omega_s} \right) + R_0
\]

For convenience, in the text we refer to those two functions shortly as erf1 and erf2.

The surround suppression’s strength was expressed by defining a contrast normalization index (CNI), computed for each time window, using the following formula

\[
\text{CNI}(t) = \frac{C_{50}(t) - C_{90}(t)}{C_{50}(t) + C_{90}(t)}
\]

where C_{50} and C_{90} are the half-saturation contrast values obtained with and without surround, respectively, for a given time window.

Positive values of the CNI indicate depression of the contrast gain...
function resulting from the surrounding stimulus, whereas negative values indicate enhancement.

The responses’ amplitude tuning with respect to the surround relative to center orientation ($\theta$) was modeled using a Gaussian tuning function of width $\sigma$, amplitude $A_o$, and offset $R_o$:

$$R(\theta) = R_o + A_o \exp\left[-(\theta - 90^\circ)^2/2\sigma^2\right]$$

(see Fig. 5B). To measure the strength of the suppression’s orientation tuning we computed a suppression index (SI) for both iso- ($SI_{iso}$) and cross-oriented ($SI_{cross}$) surrounds

$$SI_{iso} = 1 - \frac{R_o}{R_o} \quad \text{and} \quad SI_{cross} = 1 - \frac{A_o + R_o}{R_o}$$

where $R_o$ and $A_o$ are best-fit parameters of the response tuning function and $R_o$ is the response amplitude for the center-only condition. In the absence of surround modulation ($R_o = R_o; A_o + R_o = R_o$), those indices are equal to zero, whereas they are equal to 1 when either the iso- or the cross-oriented surround totally suppresses the response (as $R_o = 0$ or $A_o + R_o = 0$, respectively). Finally, we computed an orientation modulation index (OMI) as $A_o/(R_o - R_o)$. When OMI is equal to zero, surround suppression is not tuned for orientation. As OMI increases ($<1$), the surround suppression is increasingly tuned for orientation.

**Results**

Our goal was to define a limited set of descriptive, time-varying functions describing spatial integration of motion, by measuring the magnitude of the initial, open-loop phase of ocular following (Miles and Kawano 1987) at different points in time. We used sine-wave gratings of low spatial and temporal frequencies to probe spatial summation and center-surround interactions within one single spatiotemporal channel.

Spatial summation area of ocular following eye movements

First, we accurately measured the spatial summation area of the motion mechanism driving ocular following responses to a single drifting grating. Figure 1A plots, for one subject, the horizontal eye velocity profiles of tracking responses to a rightward drifting grating presented behind a circular aperture of various sizes (diameters: 3.6–43.2°). As diameter increased, the initial eye velocity rose very rapidly. Latencies increased only weakly if at all for decreasing stimulus sizes. We estimated the latency using an objective method (Krauzlis and Miles 1996) from the mean velocity profiles. Across subjects and grating motion directions, we found that mean (±SD) response latencies were of 97 ± 11 and 90 ± 8 ms for stimulus diameters of 3.6 and 7.2°, respectively. For stimuli >7° diameter, the response latencies remained largely constant at about 85 ms after stimulus motion onset (i.e., mean ± SD, 81 ± 4 ms for a stimulus diameter of 43.2°). Thus a small (nearly 15 ms) but significant increase in latency was observed only with the smallest aperture size [3.6 vs. 43.2°, t(14) = 3.14, P < 0.01; 7.2 vs. 43.2°, t(14) = 1.8, NS]. The relationship between stimulus size and amplitude of ocular responses showed a complex temporal dynamics. For the earliest part of the response (i.e., <120 ms after stimulus onset), the initial eye velocity first increased quasi-linearly up to nearly 20°, and then saturated for larger grating diameters (Fig. 1A). For the later part of the responses (>120 ms after stimulus onset) the saturation observed with the larger stimuli disappeared, so that eye velocity increased with stimulus sizes >20°, albeit with a lower rate.

![Figure 1](http://jn.physiology.org/)

**FIG. 1.** Spatial summation area for ocular following. A: eye velocity profiles of responses to a low spatial frequency vertical grating, drifting rightward behind apertures of various diameters. B: mean (±SE) change in eye position over the 95- to 115-ms time window (shaded area in A), as a function of stimulus diameter. Continuous line shows the best fit of the response amplitudes, obtained with an integral of a Gaussian spatial summation function (erfl) (subject GM). For comparison, the broken line shows the best fit obtained with the difference of 2 integral of a Gaussian functions (erf2). C: mean (±SD) across subjects and directions best-fit parameter ($\omega$) coding for the width of the excitatory spatial summation area (see METHODS for details) is plotted vs. 4 different 20-ms time windows to illustrate how temporal integration led to an increase of the spatial summation area over time. Inset: earliest (gray curve) and latest (black curve) Gaussian functions yielding the best-fit erfl functions (response amplitude vs. distance in the visual space).
The dependency of the ocular following responses on stimulus diameter at early and late times can be better seen in Fig. 1B. The data were well fitted (solid line) by a spatial summation modeled with an integral of a Gaussian function (erf1; see METHODS). However, because in one subject (GM) the initial responses displayed a small hypersaturation for large stimulus values, we also tested whether the data might not be better fitted by assuming the presence of a weak peripheral suppression by the isodirectional surround stimulus, modeled by a difference between two integral of a Gaussian functions (erf2; see METHODS). We found that the goodness of the fit was very similar to that obtained with an integral of one Gaussian model alone (i.e., $\chi^2$ were similar: 0.21 and 0.29, respectively). The optimal size (i.e., the width of the excitatory erf function) was similar for the two fits. Suppression at larger diameters was significant but weak (Subject GM, mean ± SE across motion directions: 12 ± 5.2%) and the width of the inhibitory erf function was very large. For the other subjects (as well as for GM at later times), we found only very little evidence for peripheral suppression with large stimuli. Altogether, the erf2 model did not significantly improve description of the data. Although we cannot completely rule out a small isodirectional surround suppression at very early times, our data rather suggest that, at least for the low spatial frequency gratings used (0.28 cpd), the oculomotor driving motion signal is read out by a spatial summation function that integrates isodirectional motion signals over a restricted area of the visual field. Furthermore, this central integration area grows with time.

To quantify this temporal dynamics, the extension of the excitatory spatial summation area driving ocular following was estimated by the width parameter ($\omega$) of the best-fit erf1 functions, for four different time windows (Fig. 1C). With a 0.28-cpd grating presented at full contrast, best-fit $\omega$ for the earliest time window (95–115 ms) ranged from 9.6 to 18.1° across subjects and directions, with a mean ± SD of 12.2 ± 3.2°. As can be seen, $\omega$ significantly increased over time, up to a mean value of 28.5 ± 9.2° for the latest time window [95–115 vs. 155–175: $t$-test, $t(10) = 4.09, P < 0.01$].

Effect of contrast on the optimal spatial summation area

It was previously shown at the single-neuron level that the spatial summation area of macaque V1 and MT neurons changes at very low contrast (Cavanaugh et al. 2002; Kapadia et al. 1999; Pack et al. 2005; Sceniak et al. 1999). These neuronal effects should lead to a change in ocular following responses in primates. Therefore we investigated the effects of stimulus contrast on the ocular spatial summation function. To do this, we presented the same set of grating diameters as above, but at three different contrast levels (5, 20, and 80%), selected to cover the complete contrast dynamics of ocular following (Masson and Castet 2002).

Figure 2A illustrates, for each subject, the relationships between response amplitude and stimulus diameter, observed at different contrasts. At very low contrast, all curves became more sluggish, showing only very little saturation for at least two subjects, suggesting a larger spatial summation area compared with that underlying the processing of medium and high contrast stimuli, which displayed response functions that were much steeper and saturated for stimulus diameters of about

![FIG. 2. Effect of contrast on the spatial summation area. A: for 3 subjects, the mean change in eye position (95- to 115-ms time window) is plotted as a function of center diameter for a 5% (black symbols), 20% (gray symbols), and 80% (white symbols) center contrast. Continuous lines are best fits of an integral of a Gaussian (erf1) spatial summation model. B: relationship between contrast and the best-fit estimate of the width ($\omega$) of the spatial summation function. Open circles are data obtained for each subject and direction. Crosses represent their mean (±SE). Inset: mean best-fit Gaussian functions, as response amplitude vs. distance in the visual space, for each different contrast level.](http://jn.physiology.org/)

To quantify this temporal dynamics, the extension of the excitatory spatial summation area driving ocular following was estimated by the width parameter ($\omega$) of the best-fit erf1 functions, for four different time windows (Fig. 1C). With a 0.28-cpd grating presented at full contrast, best-fit $\omega$ for the earliest time window (95–115 ms) ranged from 9.6 to 18.1° across subjects and directions, with a mean ± SD of 12.2 ± 3.2°. As can be seen, $\omega$ significantly increased over time, up to a mean value of 28.5 ± 9.2° for the latest time window [95–115 vs. 155–175: $t$-test, $t(10) = 4.09, P < 0.01$].

Effect of contrast on the optimal spatial summation area

It was previously shown at the single-neuron level that the spatial summation area of macaque V1 and MT neurons changes at very low contrast (Cavanaugh et al. 2002; Kapadia et al. 1999; Pack et al. 2005; Sceniak et al. 1999). These neuronal effects should lead to a change in ocular following responses in primates. Therefore we investigated the effects of stimulus contrast on the ocular spatial summation function. To do this, we presented the same set of grating diameters as above, but at three different contrast levels (5, 20, and 80%), selected to cover the complete contrast dynamics of ocular following (Masson and Castet 2002).

Figure 2A illustrates, for each subject, the relationships between response amplitude and stimulus diameter, observed at different contrasts. At very low contrast, all curves became more sluggish, showing only very little saturation for at least two subjects, suggesting a larger spatial summation area compared with that underlying the processing of medium and high contrast stimuli, which displayed response functions that were much steeper and saturated for stimulus diameters of about
As before, an erf function was fitted to each data set and the best-fit width parameter \( \omega \) is plotted against contrast in Fig. 2B. Increasing the contrast from 5 to 80% resulted in a significant decrease in the mean \( \omega \) from 21.6 ± 7.5 to 12.1 ± 3.3° (\( t \)-test, \( t(10) = 2.98, P < 0.01 \)), indicating that the highest-contrast gratings were sampled with spatial summation areas of about half the size of those with which the lowest-contrast ones were sampled. This reduction is further illustrated in the inset of Fig. 2B, where we plot the profiles of the mean, best-fit Gaussian envelopes. Overall, these results suggest that increasing the contrast both enhanced the response amplitude and reduced the width of the spatial integration area.

Center-surround interactions for ocular tracking: contrast normalization

Our next goal was to probe the oculomotor effects of center-surround interactions for nonisodirectional motion signals. We first found that a static peripheral grating had no effect on the tracking responses, showing that center-surround interactions occur only between two dynamic stimuli. Next, we used moving peripheral stimuli. However, recording responses to stimuli made of different motion directions presented simultaneously is inadequate because each visual motion would drive antagonistic (or synergistic depending on their relative directions) eye movements (see Supplemental Data1), hampering our objective to measure the behavioral consequences of center-surround modulations at the level of visual processing, rather than at the motor output level. To eliminate such motor interactions, we thus used a dynamic surround constructed by summing two gratings of identical orientation and spatial frequency but drifting in opposite directions at the same temporal frequency (speed). Such summation generates a “counterphase”-modulated grating that contains no net motion signal but drives neuronal mechanisms tuned for each of the two opposite motions (Qian and Andersen 1995). Indeed, when the surround was presented alone, no or only very small ocular following was observed, which was not consistent among subjects and thus most likely resulted from a small asymmetry in the individual sensitivity to the two antagonistic motion signals (Fig. 3, C and D, gray symbols). In other words, the motor effects of the two opposite peripheral motion signals cancel out. Figure 3 shows, for one subject, the dependency of ocular following on the contrast of a central grating presented behind a circular aperture of optimal diameter (20°) with or without such a counterphase grating in the surround. In Fig. 3A, we plot the mean velocity profiles of responses to a center grating presented alone.

The initial velocity of ocular following increased with grating contrast, up to a maximum reached for contrast of roughly 30–40%. Figure 3B illustrates the mean eye velocity profiles for the same sample of center contrast but now presented together with the high-contrast (80%) surround stimulus. Comparison between Fig. 3A and Fig. 3B shows that the earliest phase of ocular following responses was largely unaffected by a peripheral counterphase grating of same orientation; latencies remained largely unchanged at about 85 ms, particularly at high contrast, and initial eye velocity was only weakly af-

---

1 The Supplementary Material for this article is available online at http://jn.physiology.org/cgi/content/full/00112.2006/DC1.

---

**FIG. 3.** Center-surround interactions for early and late ocular following. A: eye velocity profiles of ocular following responses to a vertical grating drifting leftward behind a 20° diameter aperture and presented at different contrasts. B: same center grating presented with a counterphase surround grating of same orientation. Light and dark gray shaded bars indicate the early (95- to 115-ms) and late (155- to 175-ms) time windows used for measurements. C: early (95- to 115-ms time window) change in eye position plotted vs. the center contrast, for center-only (open symbols) and center-surround (closed symbols) conditions. D: same as C, but for late time window (155–175 ms). Continuous lines are best fits of the Naka–Rushton functions. Gray symbol in C and D plots the response to the surround-only condition, showing that no significant ocular following response was observed even at high contrast.
fected. On the contrary, the later part of the tracking response was strongly suppressed by the peripheral grating. In particular, the maximum eye velocity reached at the time of the closing of the oculomotor loop (about 170–180 ms) was largely reduced.

These results are seen more clearly in Fig. 3, C and D, where open and closed symbols plot the mean changes in eye position for center-alone and center-surround conditions, respectively. They were measured over two different time windows: one early (95–115 ms), centered at response initiation, and one late (155–175 ms), set immediately before the closing of the oculomotor loop. Ocular contrast-response functions were fitted with a Naka–Rushton function, best describing the contrast dynamics of oculomotor following (Masson and Castet 2002; Shehata et al. 2005). Again, the figure shows that the earliest contrast-response functions were very similar for the two conditions (Fig. 3C), whereas the two response functions largely differed for the late phase of oculomotor following (Fig. 3D). Similar results were obtained for all four subjects and for the two center grating motion directions.

**Time course of contrast dynamics**

Next, we compared early and late responses, separately for both conditions. Without surround, a clear shift toward low contrast values was observed in the late contrast-response function with respect to the early one. In the presence of a dynamic surround, on the other hand, this shift was much smaller, the early and late contrast-response functions being very similar, as can be seen by comparing Fig. 3C and Fig. 3D. To ease the comparison, in Fig. 4A we plot the half-saturation contrast ($C_{50}$) obtained from best-fit Naka–Rushton functions for the center-only versus the center-surround condition, for each subject and motion direction.

Open symbols are data obtained with the earliest time window: they were clustered only slightly above the identity line (first from bottom), indicating that, at early times, $C_{50}$ values with and without surround were very similar. Filled symbols are data obtained with the late time window. Squares plot mean values across subjects and directions. The lefward shift of the open versus the filled symbols reflects the $C_{50}$ decrease observed at late time with center-only gratings (Fig. 3D), indicating an increased contrast gain. On the other hand, we observed no decrease of the late $C_{50}$ obtained with a center-surround stimulus: rather, the data were slightly shifted in the upward direction, indicating, if anything, a decrease in contrast gain that, however, was found not to be significant (see following text). We quantified these changes by comparing $C_{50}$ values across subjects. We first compared these values across time windows, within each condition. With a center grating alone, the mean ($\pm$SD) $C_{50}$ values were of 10.2 $\pm$ 3.4 and 5.4 $\pm$ 1.2% for the early and late time windows, respectively, yielding a significant decrease over time ($t(14) = 3.78$, $P < 0.01$). With an iso-oriented surround, mean ($\pm$SD) early versus late $C_{50}$ values were not significantly different [15.6 $\pm$ 6.4 and 23.14 $\pm$ 11.8%, respectively; $t(14) = 1.58$, $P = 0.13$]. Second, we compared $C_{50}$ across center-only and center-surround conditions, for each time window. For the earliest time window, values of $C_{50}$ were not significantly different between the two conditions [$t(14) = 2.08$, $P > 0.05$], whereas in the later time window the presence of a dynamic surround significantly increased $C_{50}$ [$t(14) = 4.21$, $P < 0.01$]. The change in contrast gain was not only the result of a change in half-saturation contrast ($C_{50}$). Indeed, the presence of a peripheral surround also significantly reduced the slope ($n$) of the contrast-response function from 1.96 $\pm$ 0.36 to 1.47 $\pm$ 0.33 [mean across four subjects and two directions, $t(14) = 2.84$, $P < 0.05$], whereas the peak of the response ($R_{\text{max}}$) was not significantly affected [mean across subjects and directions, 0.026 $\pm$ 0.017 and 0.022 $\pm$ 0.015; $t(14) = 0.62$, NS].

To illustrate the temporal dynamics of the surround suppression, we computed a contrast normalization index (CNI) for each time window. The mean ($\pm$SD) index is plotted against time in Fig. 4B, showing that no significant effects of the surround were seen before about 110 ms. Indeed, with the 95- to 115-ms time windows, the mean CNI was only of 0.15 $\pm$ 0.11. However, this index rapidly increased to reach about 0.6 at roughly 140 ms after motion onset. At the end of the open-loop period, the CNI had a value of 0.61 $\pm$ 0.1: at this point in time, the surround had suppressed the center $C_{50}$ by nearly 60%. Taken together, these results indicate that the iso-oriented dynamic surround prevented the temporal evolu-
tion of the contrast dynamics of ocular following, clamping it to its earliest, quasi-linear range.

**Late surround suppression: orientation selectivity**

In the previous experiment, contrast and orientation of the surround were fixed. To investigate whether surround suppression is tuned for orientation and whether it scales with contrast, we kept both contrast (30%) and orientation (vertical) of the center grating constant, while we varied both contrast and orientation of the peripheral counterphase grating. Figure 5A illustrates, for one naïve subject, the mean eye velocity profiles of the ocular responses to a center grating presented alone (continuous lines) or together with a surround of 0, 30, 60, or 90° orientation difference with respect to the center orientation (broken lines). Left- and right-hand plots illustrate the effects of either low- or high-contrast surrounds, respectively. With low-contrast surrounds, suppression by the peripheral stimulus was nearly absent or very weak. On the contrary, high-contrast surrounds produced a strong, albeit delayed suppression. Moreover, suppression developed over time. As in the previous experiment, the first 20–30 ms of the responses were indistinguishable from the center-alone condition, but at about 110 ms after stimulus onset (vertical arrow, Fig. 5A, right), a strong suppression began to emerge. Furthermore, the surround grating induced an overall reduction in eye velocity that became gradually and partly tuned for surround orientation (Fig. 5B): at the end of the open-loop period, ocular responses were clearly more suppressed by an iso-oriented (0°) than by a cross-oriented (90°) surround. Similar results to those shown in Fig. 5 were found in three other subjects (the authors).

To illustrate these four different aspects of modulation by the surround (i.e., global and orientation-tuned suppression, contrast dependency, and temporal dynamics), we computed the mean (±SE across trials) change in horizontal eye position over 20-ms time windows, covering the whole open-loop period from 95 to 175 ms after stimulus onset. Figure 5B plots, for the same naïve subject and for several contrast values, the mean change in eye position as a function of surround orientation, for the earliest time bin where a clear response could be seen (95–115 ms), as well as for the latest time bin (155–175 ms) before the closing of the oculomotor loop (i.e., 170 ms after stimulus onset, twice the response latency of about 85 ms). Continuous lines are Gaussian functions that were fitted to the data to evaluate both amplitude and width of the orientation tuning of surround suppression. As a comparison, broken
increased with surround contrast, as can be seen by comparing right
both global suppression and its orientation specificity
increased with surround contrast, as can be seen by comparing the various curves in Fig. 5B (right) to the horizontal dotted line marking the response to the center-only condition. To quantify this result, we computed a suppression index, separately for each one of these two conditions. Figure 5C plots their mean values (±SD across subjects and directions) against surround contrast. For both iso- and cross-oriented surround gratings, the suppression indices increased with contrast from 0.07 ± 0.04 to 0.49 ± 0.12 and from 0.02 ± 0.01 to 0.33 ± 0.09, respectively. At low and medium contrast, the suppression indices of the two conditions were not significantly different. At high surround contrast, however, iso-oriented suppression was roughly 30% stronger than cross-oriented suppression [t(9) = 2.61, P < 0.05].

Increasing the surround contrast thus resulted in a stronger surround orientation tuning of the response amplitude which was best described by a Gaussian function peaking at 90° relative orientation. More specifically, we found that the modulation amplitude increased with surround contrast, whereas the width of the tuning function remained fairly constant. This can be seen in Fig. 5D, where we plot the individual best-fit orientation widths (closed symbols) as well as the mean (±SD) across subjects and directions, as a function of surround contrast: Increasing the contrast from 10 to 80% produced a small but nonsignificant widening of the orientation tuning, the mean σ of the tuning function increasing only marginally from 22.8 ± 4.5 to 26.8 ± 6.9°. However, the modulation amplitude A0 increased regularly—and significantly—with contrast, from 0.09 ± 0.05 to 0.25 ± 0.1° [t(10) = 3.5 P < 0.01; Fig. 5E]. To further illustrate the orientation tuning of surround suppression, we plotted (Fig. 5F) the mean Gaussian tuning functions drawn from the data illustrated in Fig. 5, D and E.

Finally, Fig. 6A illustrates the temporal dynamics of surround suppression and its orientation tuning by plotting best-fit tuning functions computed every 10 ms, for a rightward motion, showing that both orientation-tuned and global suppression gradually developed over time. Figure 6B plots the mean OMI (±SE across subjects and directions) against time, showing quantitatively that orientation-tuned suppression started only after 100 ms after stimulus onset, regularly increasing thereafter over the whole open-loop period: At the end of the open-loop period, about 30% of the surround suppression can be attributed to a mechanism tuned for orientation of the peripheral grating (Fig. 6B).

**Surround suppression: effect of spatial scale**

The results described above show a strong suppression of the center-driven responses by iso-oriented peripheral stimuli. Moreover, this suppression appears only about 20 ms after response onset and develops over time. Such suppression can be described as a mainly divisive change in the contrast gain, similar to the center-surround effects found in single neurons at the early stages of a visual motion stream (see Carandini 2004). However, the very large center stimuli (20°) used here are most likely to cover the receptive fields of a large sample of neurons located at different eccentricities and having different spatial summation functions. Moreover, at least in the macaque, ocular following responses depend strongly on the neuronal activity at the level of areas MT/MST (see Kawano 1999), where most neurons have receptive field sizes that are much smaller than the size of our stimulus. Using large stimuli, we might thus have driven only a small fraction of those neuronal center-surround interactions because neurons with a counterdirectional center-surround suppression mechanism would be suppressed only if located at or near the edge of our center stimuli, eventually leading to an underestimation of their impact on the ocular tracking responses. Alternatively, it is possible that the detected tracking suppression does not result from center-surround suppression at the level of each individual unit, but rather results from inhibitory interactions between two subpopulations of neurons located in the central and the peripheral part of the visual field, respectively.

Following this reasoning, we decided to test the spatial scale of surround suppression for ocular following and designed two control experiments. In a first test, we measured the contrast-response functions of eye responses to a small central patch
(10° diameter, i.e., smaller than the optimal size) presented either alone or together with an iso-oriented, counterphase grating in the surround. As a second test, we presented a set of nine small patches (each 6° in diameter) forming an array covering the same visual surface as the large stimuli used above. Each patch was presented either as center-alone or together with an iso-oriented surround. As for the case of the large stimulus used before, to measure center-surround interactions we varied the contrast of the central patch while keeping the contrast of the surround constant at 80%.

Figure 7, A and B illustrates the results obtained with an array of small patches and with a single, midsize (center diameter, 10°) patch, respectively. To maintain a sufficient number of grating cycles (≥2.5) inside each aperture, we here used a spatial frequency (0.45 cpd) different from that in the preceding experiments. To preserve the same temporal frequency (10 Hz), drifting grating speed was set to about 22%. Clearly, because we used both suboptimal stimulus sizes and spatiotemporal properties, ocular following responses were very small and somewhat delayed (about 5–10 ms). The five 20-ms time windows were adjusted accordingly, starting at 85 ms after stimulus onset. Open and closed symbols plot the amplitude of ocular following responses over three of these time windows, two early (85–105 and 105–125 ms, left) and one late (165–185 ms, right) for center-alone and center-surround conditions, respectively. Continuous lines are best-fit Naka–Rushton functions. As found with a single, large (20° diameter) patch, the effect of the added surround developed in time. Such suppressive effect resulted in a rightward shift of the contrast-response function, yielding a large and significant increase in the best-fit $C_{50}$ parameters with respect to the center-only condition, for the later phase of the responses (time window: 165–185 ms). On the contrary, during the earliest phase (time-window: 85–105 ms) of ocular following the responses were only weakly, if at all, affected by the presence of an iso-oriented surround. A closer examination of the different contrast-response functions reveals that, as observed for the large stimuli used above, ocular contrast-response functions for the center-only condition gradually shifted to lower contrast values over time, whereas those obtained in the presence of a surround remained remarkably constant. Such dynamics was very similar for both types of motion stimuli (single and multiple patches). Finally, in both experiments, the maximum amplitudes of the ocular responses were substantially reduced by the presence of a surround, a result that is somewhat different from that reported above for a large stimulus. This reduction in the ocular response gain might be attributable either to the slightly different spatial frequency and/or speed used or to stronger surround suppression. These two hypotheses remain to be further investigated.

The above effects are summarized in Fig. 8, A and B, where we plot the half-saturation contrast values measured with center-alone and center-surface conditions one against the other, for three subjects and each grating motion direction. Note that, because of the small stimulus sizes used here, the early responses were very small (see first column of Fig. 7), yielding only noisy estimates of change in eye position over the first 20 ms and thus poorer fits for this time window. For this reason, in Fig. 8 we plotted the data obtained from three different time windows. As expected, the $C_{50}$ values for the
center stimulus decreased between the first (85–105 ms) and the second (105–125 ms), as well as between the second and the third (165–185 ms) time window (arrows in Fig. 8A and B). Moreover, $C_{50}$ for the center-surround condition remained roughly constant for the micropattern (Fig. 8B). For the single patch (the smallest stimulus in terms of overall surface), on the other hand, $C_{50}$ decreased between the first and the second time window even for the center-surround condition (first arrow segment, Fig. 8A), as opposed to the data obtained with a large stimulus (Fig. 4A). This apparent discrepancy can be ascribed to a poor fit of the Naka–Rushton functions arising from the low SNR of the measurement at the earliest times. Indeed, between the second and the latest time window, where the quality of the fit was much better, the $C_{50}$ values for the center-surround stimulus again increased (second arrow segment in Fig. 8A), as found with the large-size stimulus used above (Fig. 4A).

With a small ($10^\circ$ diameter) central grating (Fig. 8A), the earliest $C_{50}$ points (open circle) were scattered around a mean of $19.1 \pm 3.5\%$ (center-alone) and $32.8 \pm 4.5\%$ (center-surround) (open square). In the next time window (gray circles), the mean $C_{50}$ point was found at $6.7 \pm 1.7$ and $11 \pm 3.5\%$, indicating a small (roughly 4%) albeit significant $t$-test, $t(10) = 2.7, P = 0.02$ difference in mean $C_{50}$ values. At the end of the open-loop period, data points were scattered around mean values of $4.4 \pm 0.8\%$ (center-alone) and $18.1 \pm 4.1\%$ (center-surround), indicating a threefold difference in half-saturation contrast $t$-test, $t(10) = 8.03, P < 0.0001$. Mean $C_{50}$ values obtained with the center-only condition at the earliest time window and the center-surround condition at the latest time window were not significantly different [19.1 ± 3.5 and 18.1 ± 4.1%, respectively, $t$-test, $t(10) = 0.8, P > 0.4$]. With the micropattern (Fig. 8B), the earliest $C_{50}$ data points were scattered around mean values of $28.9 \pm 3.5\%$ (center-alone conditions) and $20.2 \pm 4.8\%$ (center-surround conditions). In the 105- to 125-ms time window, the data were scattered around a mean value of $8.8 \pm 2.4$ and $21.9 \pm 9.2\%$. Within 100 ms of tracking, data points moved purely along the horizontal axis, scattering around $4.5 \pm 0.5$ and $25.2 \pm 2.2\%$, which is a fivefold increase in mean $C_{50}$ [$t$-test, $t(10) = 22.5, P < 0.0001$]. Moreover, $C_{50}$ values obtained for the center-only condition in the earliest time window and for the center-surround condition with the last time window were not significantly different [28.9 ± 3.5 and 25.2 ± 2.2%, respectively, $t$-test, $t(10) = 2.19, P > 0.05$]. This result demonstrates that an iso-oriented surround clamps the contrast gain parameter to its initial value, as opposed to the case of center-alone conditions, where $C_{50}$ dramatically decreased over time.

Finally, Fig. 8C plots the CNI for the two control conditions (open symbols: small single patch; closed symbols: multiple small patches) together with the single large center grating condition, replotted from Fig. 4B. It is clear that, for all conditions, the surround suppression grew over time with very similar dynamics. It reached a maximum of about 50–60% after 100 ms. Mean (±SE) maximum CNI values were of $56 \pm 6$ and $69 \pm 3\%$ for single and multiple small patches conditions, respectively [for comparison, with a large ($20^\circ$ diameter) central patch, we observed a mean CNI of $61 \pm 10\%$; see above]. Notice that for the case of the micropattern stimulus, the CNI during the earliest time window was negative ($-20 \pm 0.13$), indicating that, for some subjects, $C_{50}$ values were actually smaller in the presence of a surround, meaning that...
somewhat larger ocular following responses were observed at very low center contrast in the presence of a surround. We attribute this phenomenon to a slight contamination of the responses by one of the two opposite motion direction components of the peripheral stimulus, as shown by the small but significant transient tracking response observed with the surround-stimulus alone (the gray symbol in Fig. 7A, left) for the earliest time window. Whereas in the case of large responses to the center stimulus (see Figs. 3C and 4B) this surround-only response is negligible, its effect becomes relevant (Fig. 8C) when the center responses are small (Fig. 7A).

Altogether, very similar results were obtained for center-surround stimuli of different sizes and for different spatial arrangements, suggesting that the observed center-surround effects result from neuronal interactions at the population level, rather than from center-surround properties of single neurons.

**Discussion**

By probing spatial summation and center-surround interactions underlying the rapid initiation of tracking eye movements, we have mapped several key properties of motion integration for ocular following. These properties can be described with a small set of operators that summarize both linear and nonlinear mechanisms and define a behavioral receptive field (BRF), in a way reminiscent of the neuronal and perceptual receptive fields used to synthesize an input–output transfer function at a given level of brain organization. We suggest thereafter that properties of the BRF emerge from the activity of large populations of neurons that are cascaded at different stages of the cortical motion pathway. Moreover, several functional consequences for ocular tracking can be predicted from the properties of this BRF. Finally, such definition of a BRF should contribute to a better understanding of how fundamental properties of motion-selective neurons are readout at the population level in the context of a simple action system (Priebe and Lisberger 2004; Takemura et al. 2001).

**Behavioral receptive field: spatial properties**

The earliest phase of ocular tracking is driven by a mechanism that integrates center motion signals in a quasi-linear fashion, up to an optimal stimulus diameter of about 20° for high-contrast, low spatial frequency stimuli. This optimal size defines the driving center of the BRF. Similar values have been found using low spatial frequency random-dot patterns in monkeys (Miles et al. 1986) and humans (Gellman et al. 1990). Altogether, these results demonstrate that ocular following responses are not driven by the *en masse* motion of the visual field but rather by a large central subportion of it (see Miles 1998; Miles and Kawano 1987). Within the BRF’s driving center, motion signals are linearly integrated, as shown by the linear relationship found between response amplitude and stimulus size in the 0–20° diameter range. Such linear integration is consistent with our previous finding that ocular following is initiated by a linear combination (i.e., vector sum/average) of the different motion direction signals present in the central visual field (Masson 2004; Masson and Castet 2002; Masson et al. 2001).

The BRF, however, also exhibits several nonlinearities. First, the saturation of the initial tracking responses found for stimuli >20° suggests that, at short latencies (<100 ms), neurons in the peripheral visual field are given little weight in the integration of motion signals. Second, the only marginal or absent decrease of the initial response amplitudes for stimuli exceeding the optimal size (20°) suggests that suppression by peripheral stimuli of the same motion direction as that of the center is only very weak, if not absent, with low spatial frequency inputs. Third, the suppressive effects observed in the presence of a dynamic surround unravel the suppressive effect of nonsidirectional peripheral motion signals. Indeed, the experiments performed with supraoptimal size center stimuli show that peripheral motion isodirectional with the center have no or only small effects; the suppression observed with counterphase peripheral gratings can thus be attributed only to the peripheral motion component of direction opposite to that of the center. Finally, showing that surround-only stimuli did not trigger ocular following, we ruled out any possible interactions between motor responses. Such visual peripheral suppression acts mainly by dividing the contrast gain of ocular tracking and mimics many of the properties of the divisive process postulated to explain contrast normalization at different stages of the motion pathways: a rightward shift in the contrast-response function (Figs. 3 and 7), a suppression scaled with surround contrast (Fig. 5) and partially tuned for orientation, peaking with iso-oriented surround gratings (Fig. 5) (see Heeger 1993; Schwartz and Simoncelli 2001; Simoncelli and Heeger 1998).

Finally, the last experiments demonstrate that contrast normalization for ocular tracking is not restricted to one particular, optimal spatial scale or geometry. Instead, reducing the size of the center-surround stimulus below the optimal center size or presenting several center-surround patches inside the BRF produced a similar contrast gain change together with a stronger reduction in response amplitudes. This indicates that interactions between visual signals occur at several spatial scales within the BRF.

Because we used counterphase gratings as peripheral stimuli, we could not directly distinguish between directional and orientation tuning of the surround suppression. However, oriented static surrounds have no modulatory effects and, most important, in the case of iso-oriented surrounds, suppression resulted from gratings moving in the opposite rather than in the same direction, strongly suggesting that this selective component is actually tuned for motion *direction* and that suppression results from interactions between different subpopulations of direction-selective neurons. Moreover, because orientation/direction tuning explains only a part of surround suppression, our results show that contrast normalization also contains a global suppressive component in addition to the tuned one. All these nonlinear properties have already been demonstrated for motion processing at both neuronal (e.g., Cavanaugh et al. 2002; Heuer and Britten 2002; also see Carandini 2004) and perceptual (e.g., Chubb et al. 1989; Petrov et al. 2005; Solomon et al. 1993; Xing and Heeger 2001) levels. To our knowledge, however, the results presented herein are the first demonstration of the functional consequences of these early visual nonlinearities on motor behavior.

It shall be noticed, however, that the peripheral suppression reported herein is largely different from the peripheral modulation reported earlier by Miles and co-workers for ocular following (monkeys: Miles et al. 1986; humans: Gellman et al. 1990). With random-dot patterns, they observed a facilitation...
of center-driven responses with an antagonistic peripheral motion, an effect called antiphase enhancement. Several differences in the experimental conditions can explain the apparent contradiction between the two studies. More important, in both humans and monkeys, antiphase enhancement was observed only with very large center motion (>40°), that is, for center stimuli much larger than the optimal sizes observed herein with low spatial frequency gratings. Second, this antiphase enhancement was found only very late in the response, roughly after the closing of the oculomotor loop. During the initial part of the responses, a weak suppressive influence was observed that can be attributed to a cancellation between two antagonistic responses (see also supplemental data). In fact, Masson et al. (2001) observed that when motions in opposite directions are presented in the central part of the visual field, no net ocular following responses were seen, suggesting that each motion drives a tracking response in its own direction. Further investigation of the differences and similarities between ocular following cancellation and surround suppression is clearly needed to decipher the visuomotor and visual interactions driven by complex motion stimuli.

Behavioral receptive field: temporal properties

We propose that the BRF includes a weighted linear integration and divisive contrast normalization through center-surround interactions. We probed their temporal dynamics by measuring these operators at different points in time over the nearly 90-ms open-loop period. Temporal motion integration for ocular tracking has two different signatures. First, in the absence of a suppressive surround, there is both a progressive expansion of the BRF’s center (Fig. 1) and a shift of the ocular contrast-response function toward lower contrast values (i.e., an increase in the contrast gain; Fig. 3). These changes are gradual but seem to saturate at the end of the open-loop period. A dynamic surround prevents the latter temporal evolution by both clamping the contrast gain to its earliest range and decreasing the response gain. Such suppressive interaction gradually develops over time so that about 90 ms after response onset the contrast gain decreased by nearly 60% with respect to its value observed in the center-only condition. At this time, the maximum eye velocity obtained with the highest contrast is also lowered, indicating an overall reduction of the ocular tracking response.

There is a second striking aspect of the BRF’s temporal properties. Suppression by surround motion of direction different from that of the center starts only about 20 ms after response onset, suggesting that surround signals have an effect on the center-driven responses only after a fixed delay. We previously showed that linear integration of local grating motion information is done at ultrashort latency (about 80–85 ms) (Masson and Castet 2002). Moreover, inverted ocular following responses to reversed-phi motion have the same latency as that of normal responses (Masson et al. 2002). Ocular contrast gain is also set immediately at response onset (Masson and Castet 2002; Sheliga et al. 2005), albeit its value changes over time. On the contrary, nonlinear motion computations are delayed by about 20 ms, as shown by response timing to pattern motion direction when using plaids (Masson et al. 2002). The similarity between the timing found here for center-surround interactions and that observed for one-dimensional/two-dimensional (1D–2D) motion integration using plaids (Masson and Castet 2002) is of particular interest in the context of recent modeling work, which has demonstrated that both nonlinear contrast gain control and inhibitory inputs are necessary to reconstruct pattern motion direction (Rust et al. 2005), leading to the speculation that the temporal dynamics of 2D motion integration might in fact reflect the temporal dynamics of these two nonlinear elements.

A neuronal population coding for the BRF

Characteristics of the BRF should be seen as features emerging from the properties of neuronal populations involved in integrating motion signals for driving eye movements. Short-latency ocular following responses were first documented in monkeys (Miles et al. 1986) and numerous studies have been conducted since to elucidate their neuronal substrate, particularly the key roles played by areas MT and MST (see Kawano 1999). First, correlated neuronal activity was found in areas MT and MST, preceding ocular following onset by about 10 ms (Kawano et al. 1994). Second, lesions of area MST completely abolish ocular following (Takemura et al. 2002). Third, another type of short-latency ocular responses, vergence, is driven by a vector averaging readout of MT/MST disparity-selective neurons (Takemura et al. 2001). Similar studies have yet to be conducted for ocular following, although these physiological findings are consistent with the idea that initiation of smooth pursuit eye movement is driven by a linear readout (vector sum/averaging) of a population of MT neurons tuned for direction and speed (Pribe and Lisberger 2004). Finally, it has been suggested that the dynamics of ocular following responses to plaids (Masson and Castet 2002) closely mimics the time course of neuronal selectivity for pattern motion direction in area MT (Pack et al. 2001; Smith et al. 2005).

What are the neuronal population mechanisms implementing a BRF? Most properties of ocular following are very similar in both humans and monkeys (Bussettini et al. 1996) and one might therefore speculate that properties of the BRF can be explained from those observed at neuronal levels in monkeys. First, the ocular following spatial summation function can be explained by a simple model, where activity is linearly integrated over a population of MT neurons panning the visual field, with receptive field sizes increasing linearly with eccentricity (Albright and Desimone 1987). The fact that we found only little if any evidence for isodirectional surround suppression at all spatial scales down to <5° suggests that information at low spatial frequency might be linearly integrated over a neuronal population. A suitable candidate could be wide field neurons, a subclass of MT neurons having large receptive field sizes and weak center-surround suppression (Born 2000; Born and Tootell 1991). Interestingly, wide-field neurons have been found to be involved in global motion processing and initiation of reflexive tracking eye movements in monkeys (Born et al. 2000).

A simple way to account for the observed saturation for stimulus diameters >20° consists then in assuming that the activity of the individual neurons is weighted at the readout level when integrating over the visual field, the weighting function being characterized by a sharp peripheral decay so that the contribution of peripheral neurons rapidly vanishes.
beyond a given eccentricity. The dependency of spatial summation on contrast and temporal integration can then be easily explained within such a linear weighted integration. First, the spatial integration’s peripheral cutoff could be shifted toward larger values simply by an increase in neuronal receptive field sizes for low contrast, as observed in both V1 (Cavanaugh et al. 2002; Kapadia et al. 1999; Sceniak et al. 1999) and MT (Pack et al. 2005). Second, a broadening of the weight function itself with time would allow increasingly peripheral neurons to contribute to the motion signal readout.

How to explain center-surround interactions? The simplest hypothesis of neurons having a large excitatory receptive field (essentially matching the stimulus) and a suppressive surround activated by the nonisodirectional motion in the periphery is unlikely to be true because we show similar center-surround effects using both small and large stimuli as well as single- and multipatch patterns. Furthermore, neurons with such large excitatory fields in macaque areas MT and MST have little surround suppression (Born 2000). On the contrary, neurons with a strong center-surround interaction tend to have much smaller receptive fields (Born and Tootell 1991; Eifuku and Wurtz 1998; Pack et al. 2005). With large center stimuli, most of these units would be suppressed. Moreover, with large center-surround stimuli, neurons located only at the center-surround border would be modulated, resulting in only weak contrast gain change. Thus the strong modulation detected here, similar for large, small, and micropattern stimuli, indicates that the BRF is unlikely to result simply from the properties of neurons that belong to a particular subpopulation of direction-selective neurons in MT/MST.

We propose that the BRF should be understood at a more functional level as the product of interactions between different neuronal populations activated by the two motion stimuli (center vs. periphery). Similarly to what has been proposed at the neuronal level, one can assume that the two populations form an excitatory (or driving) and a suppressive (or modulating) field, respectively (see Carandini 2004). The suppressive field is formed by a subpopulation of neurons with direction selectivity different from that of the center motion direction. This suppressive population implements contrast normalization by dividing the activity of the neuronal population detecting center motion (Simoncelli and Heeger 1998) and units tuned for opposite motion direction have the highest weight in the divisive term (Schwartz and Simoncelli 2001). The structure of the BRF could then be modeled as a ratio of the two populations forming the excitatory and suppressive fields, somewhat similar to the neuronal ratio of a Gaussian model (Bonin et al. 2005; Carandini 2004; Cavanaugh et al. 2002). Such a mechanism is independent of the spatial scale of center-surround stimuli. Moreover, if the suppressive population’s weight decreases with eccentricity, the observed stronger suppression with multipatch stimuli is expected because with these stimuli, surrounding motions covered a larger portion of the central visual field.

The neuronal implementation of the BRF raises several questions. Are the suppressive interactions implemented at the readout stage or do they reflect local neuronal interactions? Temporal properties of surround suppression may provide a hint to decide between the two possibilities. Indeed, the observed delayed suppression is inconsistent with interactions at the readout stage between the driving and suppressive populations. On the contrary, there is some evidence that local inhibitory interactions such as revealed by neuronal center-surround structures are delayed relative to center-driven responses: in macaque area V1, surround modulation has been found to be delayed relative to the center-driven response by about 15–20 ms (Knierim and Van Essen 1992) or about 50 ms (Lamme 1995). The time course of peripheral suppression in macaque area V1 was carefully reexamined by Bair et al. (2003), who found an average roughly 10-ms latency difference between onsets for central excitation and peripheral suppression. In MT neurons, surround responses are delayed by about 15 ms relative to center-driven response onset (Perge et al. 2005). This delayed surround input might explain why ocular surround suppression is delayed by about 20 ms relative to tracking onset. Finally, implementation of surround suppression by local interactions between populations tuned for different directions could explain why multipatch stimuli produce a somewhat stronger suppression than that of single patches: Although the total surround surface was roughly equal between the two conditions, micropattern displays increase the probability of activating local, short-range inhibitory connections between nearby neurons of different direction selectivity.

Functional consequences of the dynamics of motion integration

The idea of a BRF might help in understanding how visual information is used to control eye movements. Ocular following responses act to quickly stabilize the retinal images of an object of interest, foverted by the preceding saccade. In this respect, it is crucial to restrict motion analysis to this object and ignore other competing motions from the surround. By emphasizing central visual information, the BRF center helps such automatic object motion segmentation (Miles and Kawano 1987). Widening the driving part of the BRF at low contrast is consistent with earlier psychophysical studies showing that increasing the size of a low-contrast moving stimulus enhances its visibility (Anderson and Burr 1991; Watson and Turano 1995). Such a mechanism is postulated to increase the signal-to-noise ratio of neural processing and therefore to improve the accuracy of speed estimates (Verghese and Stone 1996), a critical signal for driving tracking eye movements.

What could be the role of peripheral suppression? Center-surround interactions are involved in several fundamental mechanisms of motion processing (see Albrecht et al. 2003). One of them is to set the dynamic range of neuronal responses through contrast normalization (Albrecht and Hamilton 1982; Heeger 1992; Simoncelli and Heeger 1998). This has important advantages in the context of smooth eye movements. First, divisive normalization can act to optimally set the gain of the sensorimotor transformation by using visual information from the surround rather than waiting for visual feedback at the closing of the oculomotor loop. Interestingly, it was previously found that the gain of initial smooth pursuit tends to decrease in the presence of a visual surround (monkeys: Keller and Kahn 1986; humans: Kimmig et al. 1992). A second advantage of controlling ocular contrast gain would be to stabilize oculomotor performance against sudden lighting changes, as neuronal contrast gain control does for early visual neurons (Mante et al. 2005). A third advantage would be related to another putative role of divisive normalization. Schwarz and Simon-
celli (2001) suggested that orientation-tuned divisive normalization can be used to remove statistical redundancy from the visual image. A by-product of this is to amplify other motion signals such as 2D cues. The present results thus call for further studies to investigate whether surround suppression can be used to boost the 2D motion integration dynamics for ocular following (Masson and Castet 2002). We suggest that the present results support the notion of a BRF that may open the door for future research concerning the dynamics of motion integration for the tracking of complex stimuli.

ACKNOWLEDGMENTS

We are grateful to the two naïve subjects (AM and JW) for carrying out the numerous, daily eye movement recordings. We thank E. Castet and L. Perrinet for help and advice and B. Beutter, F. Chavane, Y. Frégna, and Y. Trotter for critical reading of the manuscript.

GRANTS

This work was supported by Grant CTI-2003 from Centre National de la Recherche Scientifique and Grant ACI-JC-2001-5052 from Ministère de la Recherche. I. Vanzetta was supported by a fellowship from the Fondation pour la Recherche Médicale.

REFERENCES


Rust NC, Simoncelli EP, and Movshon JA. Neurons in MT compute pattern direction by pooling excitatory and suppressive inputs. J Vis 5: 83a, 2005


Smith MA, Majaj NJ, and Movshon JA. Dynamics of motion signaling by neurons in macaque area MT. Nat Neurosci 8: 220–228, 2005.


