Ocular-following responses to white noise stimuli in humans reveal a novel nonlinearity that results from temporal sampling

Boris M. Sheliga
Laboratory of Sensorimotor Research, National Eye Institute, National Institutes of Health, Bethesda, MD, USA

Christian Quaia
Laboratory of Sensorimotor Research, National Eye Institute, National Institutes of Health, Bethesda, MD, USA

Edmond J. FitzGibbon
Laboratory of Sensorimotor Research, National Eye Institute, National Institutes of Health, Bethesda, MD, USA

Bruce G. Cumming
Laboratory of Sensorimotor Research, National Eye Institute, National Institutes of Health, Bethesda, MD, USA

White noise stimuli are frequently used to study the visual processing of broadband images in the laboratory. A common goal is to describe how responses are derived from Fourier components in the image. We investigated this issue by recording the ocular-following responses (OFRs) to white noise stimuli in human subjects. For a given speed we compared OFRs to unfiltered white noise with those to noise filtered with band-pass filters and notch filters. Removing components with low spatial frequency (SF) reduced OFR magnitudes, and the SF associated with the greatest reduction matched the SF that produced the maximal response when presented alone. This reduction declined rapidly with SF, compatible with a winner-take-all operation. Removing higher SF components increased OFR magnitudes. For higher speeds this effect became larger and propagated toward lower SFs. All of these effects were quantitatively well described by a model that combined two factors: (a) an excitatory drive that reflected the OFRs to individual Fourier components and (b) a suppression by higher SF channels where the temporal sampling of the display led to flicker. This nonlinear interaction has an important practical implication: Even with high refresh rates (150 Hz), the temporal sampling introduced by visual displays has a significant impact on visual processing. For instance, we show that this distorts speed tuning curves, shifting the peak to lower speeds. Careful attention to spectral content, in the light of this nonlinearity, is necessary to minimize the resulting artifact when using white noise patterns undergoing apparent motion.

Introduction

One-dimensional (1D; e.g., barcode) and two-dimensional (2D; e.g., random checkerboard or dots) white noise patterns are widely used in laboratory settings to study the visual processing of images that are spatiotemporally broadband. Since neurons at the earliest stages of visual processing respond to stimuli over a limited range of spatial and temporal frequencies (for review, see De Valois & De Valois, 1988), observed responses to broadband stimuli can provide insight into how these signals are integrated in subsequent processing stages. Some of these interactions are known to be nonlinear. For instance, when two gratings are combined with different contrasts, responses are dominated by the high-contrast grating in a "winner-take-all" fashion. The effects were observed both in

Citation: Sheliga, B. M., Quaia, C., FitzGibbon, E. J., & Cumming, B. G. (2016). Ocular-following responses to white noise stimuli in humans reveal a novel nonlinearity that results from temporal sampling. *Journal of Vision, 16*(1):8, 1–22, doi:10.1167/16.1.8.

doi: 10.1167/16.1.8 Received September 8, 2015; published January 13, 2016 ISSN 1534-7362

neuronal populations (areas MT/MST; Kumbhani, Saber, Majaj, Tailby, & Movshon, 2008; Miura, Inaba, Aoki, & Kawano, 2014b) and behaviorally (Liu & Sperling, 2006; Sheliga, Kodaka, FitzGibbon, & Miles, 2006). Some nonlinearities develop during the course of visual processing: The speed tuning of most direction-selective neurons in the striate cortex can be well described by a linear summation of spatiotemporal components (Priebe, Lisberger, & Movshon, 2006), whereas in area MT the speed tuning bandwidth to noise is narrower than that predicted from grating responses (Priebe, Cassanello, & Lisberger, 2003).

In this study we quantitatively evaluated the contribution of different spatiotemporal components in generating human ocular-following responses (OFRs) to white noise stimuli. To accomplish this task we compared responses to three stimuli: white noise, band-pass filtered noise, and noise filtered with a notch filter (the complement of each band-pass filter). The OFR is a short-latency tracking eye movement evoked by motion of a textured pattern (Gellman, Carl, & Miles, 1990; Miles, Kawano, & Optican, 1986), whose characteristics (e.g., amplitude, latency) provide a behavioral signature of neuronal mechanisms operating at early stages of visual processing (for review, see Masson & Perrinet, 2012; Miles, 1998; Miles & Sheliga, 2010).

We show that the magnitude of OFRs to 1D noise stimuli can be increased by removing components with high spatial frequency (SF), suggesting that these have a suppressive effect. We show that this can be explained as a consequence of temporal sampling of the display combined with a novel nonlinearity we demonstrate: an inhibitory effect of flicker. We show that this distorts estimates of speed tuning measured with temporally sampled displays. Some preliminary results of this study were presented in abstract form elsewhere (Sheliga, Quaia, FitzGibbon, & Cumming, 2014a, 2014b).

Experiment 1

Materials and method

Many of the techniques are described only briefly because they are similar to those used in this laboratory in the past (e.g., Sheliga, Chen, FitzGibbon, & Miles, 2005). Experimental protocols were approved by the institutional review committee concerned with the use of human subjects. Our research was carried out in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki), and informed consent was obtained for experimentation with human subjects.

Subjects

Three subjects took part in this study: two were authors (BMS and EJF), and the third was a paid volunteer (AGB) who was naïve to the purpose of the experiments. All subjects had normal or corrected-to-normal vision. Viewing was binocular.

Eye-movement recording

The horizontal and vertical positions of one eye (right: BMS and EJF; left: AGB) were recorded with an electromagnetic induction technique (Robinson, 1963). A scleral search coil was embedded in a silastin ring (Collewijn, Van Der Mark, & Jansen, 1975), as described by Yang, FitzGibbon, and Miles (2003).

Visual display and stimuli

Experiments took place in a darkened room. Subjects' heads were positioned using a head band and adjustable rests (for the forehead and chin). Visual stimuli were presented on a 21-in. cathode ray tube (CRT) monitor located straight ahead 45.7 cm from the corneal vertex. The monitor screen had 1024×768 pixel resolution (20.55 pixels/°, directly ahead of the eyes), a vertical refresh rate of 150 Hz, and a mean luminance of 20.8 cd/m^2. The RGB signals from the video card reached the RGB inputs of the monitor via an attenuator (Pelli, 1997) and a video signal splitter (AC085A-R2; Black Box Corporation, Lawrence, PA), providing the 11-bit grayscale resolution of black-and-white images.

Visual stimuli were seen through a rectangular aperture (approximately $25° \times 25°$; 512×512 pixels) centered directly ahead of the eyes. Stimulus motion was horizontal at approximately 22°, 44°, or 88°/s (achieved by shifting an image by 3, 6, or 12 pixels each video frame, respectively). Three types of stimuli were implemented.

- White noise: Vertical 1D binary white noise stimuli (vertical barcode, such as the one shown in Figure 1A) were constructed by randomly assigning a "black" or "white" value to each successive column of pixels. The upper panel of Figure 1B shows the results of 1D fast Fourier transformation (FFT) of such a stimulus along the axis of motion and illustrates the fact that, on average, all SFs are equally represented. The 1D noise stimuli always had 32% root mean square (RMS) contrast.
- Band-pass filtered noise: The 1D white noise images (32% RMS contrast) were filtered using a band-pass filter that was Gaussian on a log scale. The central SF of the filter varied from 0.0625 to 4 cpd in half-octave increments, while the full width at half maximum (FWHM) was always set to two octaves. We used a

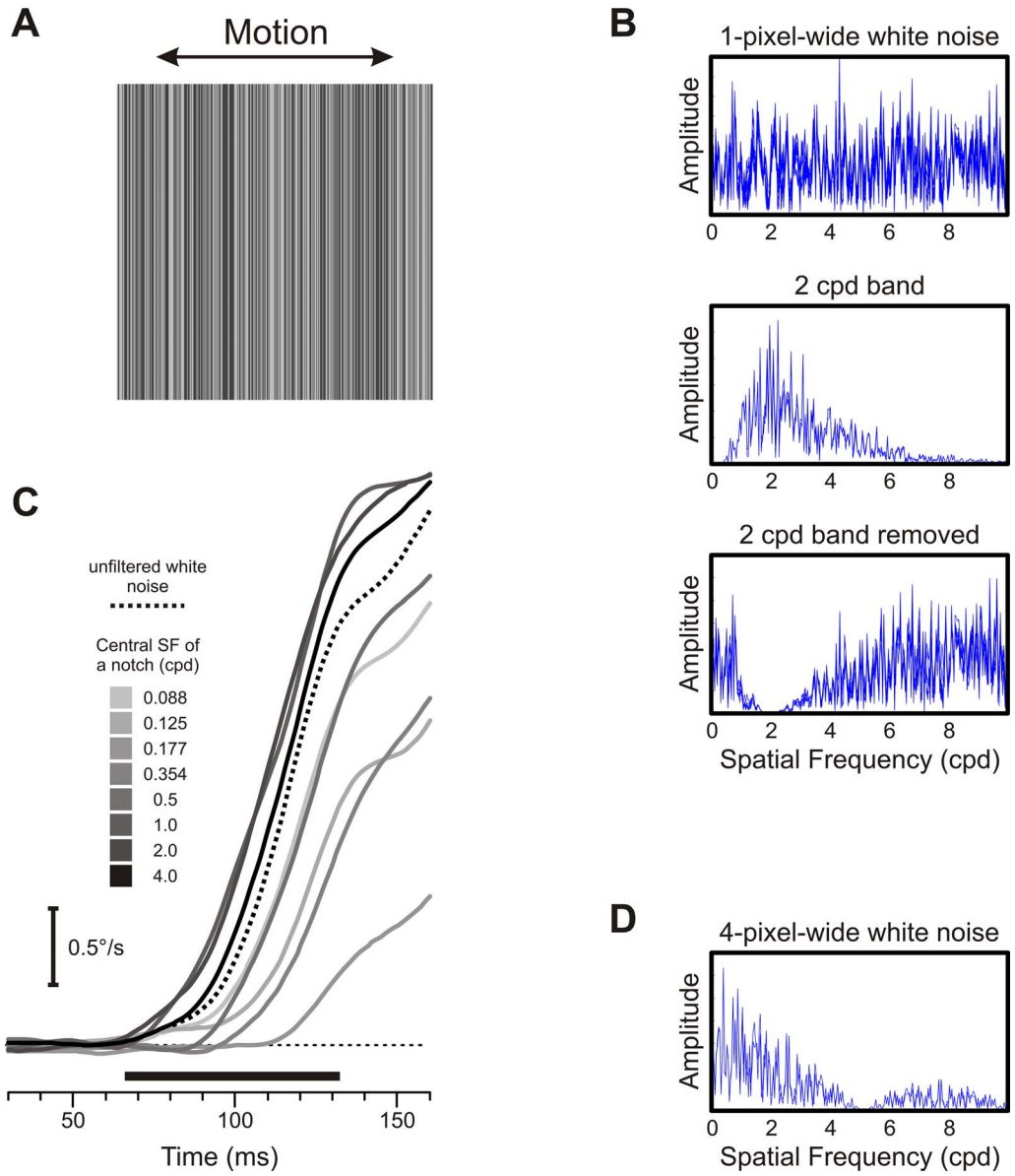

Figure 1. Experiment 1. (A) An example of vertical 1D white noise stimulus (i.e., vertical barcode): a scaled version of a 25° × 25° 1-pixel-wide pattern. (B) Upper panel: Fourier composition of a 1D 1-pixel-wide white noise stimulus. Middle panel: Fourier composition of a 2-cpd filtered white noise stimulus. Lower panel: Fourier composition of a 1D 1-pixel-wide white noise stimulus with notch at 2 cpd. (C) Mean eye velocity profiles over time to unfiltered white noise (dotted black trace) as well as to white noise stimuli depleted of various SF bands (noted by grayscale coding of velocity traces; see the insert). Subject BMS. Each trace is the mean response to 152 to 161 repetitions of the stimulus. Abscissa shows the time from the stimulus onset; horizontal dotted line represents zero velocity; horizontal thick black line beneath the traces indicates the response measurement window. (D) Fourier composition of a 1D 4-pixel-wide white noise stimulus.

fixed bandwidth on a log scale because this approximately describes the SF tuning of neurons in the striate cortex (e.g., De Valois, Albrecht, & Thorell, 1982). Although De Valois et al. (1982) showed a slight decrease in octave bandwidth with SF, the bandwidth is more nearly constant in log frequency than in linear frequency. Our two-octave bandwidth is slightly larger than the mean reported by De Valois et al. (1982), but we chose it because it produces higher RMS contrasts and hence more robust OFRs. The middle panel of Figure 1B shows the FFT of a filtered noise stimulus. Because the bandwidth was fixed on a log scale, filtered noise stimuli with higher central SF had higher RMS contrast.

- Notch filtered noise: A range of SFs were removed from the 1D noise using the same Gaussian function of SF as used for band-pass filtering. Thus, these filters were the complement of the band-pass filters:

The sum of the band-pass filtered image and the notch filtered image reconstructs the 1D binary noise pattern. The lower panel of Figure 1B displays the FFT of a stimulus with a notch at 2 cpd.

OFRs to band-pass filtered stimuli were studied for the 44°/s speed only and were randomly interleaved in a single block of trials (54 trials/block), with the other two stimulus types moving at this same speed. OFRs to stimuli moving at 22°/s and 88°/s were run in separate experimental sessions, and a single block (56 trials/block) comprised randomly interleaved unfiltered and notch filtered stimuli moving at either speed. In Experiment 1C (below), responses to band-pass stimuli were studied with a fixed temporal frequency (TF).

In Experiment 1 we did not adjust the luminance contrast of the stimuli after filtering. Because our filters had constant log bandwidth, they removed more contrast energy with increasing center frequency. If we had normalized the contrast after filtering, then these stimuli would have contained more contrast energy at low frequencies that unfiltered stimuli. This would have made any increase in OFR difficult to interpret. Without normalization, the fact that simply removing contrast energy leads to an increase in OFR clearly indicates that the components removed have a suppressive effect in the unfiltered stimulus.

Experiment 1B

Experiment 1B examined whether the OFR TF tuning depended on SF. Vertical sinusoidal gratings (32% Michelson contrast) shifted horizontally with a range of TFs: 1.67, 2.5, 4.17, 6.25, 8.33, 12.5, 16.7, 25, and 37.5 cycles/s. Two SFs were tested in each subject: 0.125 and 0.5 cpd. As in all other experiments in this article, stimuli occupied a rectangular aperture (approximately 25° × 25°) centered at fixation. A single block of trials had 36 randomly interleaved stimuli: nine TFs, two SFs, and two directions of motion (left vs. right).

Experiment 1C

Experiment 1C examined the OFR SF tuning for band-pass filtered noise stimuli. Central SFs varied from 0.0625 to 2 cpd in octave increments, and the FWHM was two octaves. The speed of each stimulus was chosen such that its central SF moved with the near-optimal TF: 23.4, 19, and 17.1 cycles/s for subjects AGB, BMS, and EJF, respectively. A single block of trials had 12 randomly interleaved stimuli: six central SFs and two directions of motion (left vs. right). Note that because of our fixed octave bandwidth, RMS contrast increased with SF. Therefore, this differs from a traditional measurement with gratings, but it does more directly measure the contribution of frequency components present in our white noise stimulus.

Experiment 1D

Experiment 1D assessed the impact of temporal incoherence ("flicker") on the OFRs. The stimulus was the sum of two band-pass filtered noise stimuli: one moving and one containing a different randomly chosen filtered noise sample on each frame (which we call *flicker*). The speed of the moving stimulus was chosen to produce a TF of 19.2 Hz for the central SF. For comparison, moving stimuli were also presented in isolation. The FWHM of all filters (Gaussian envelope on log scale) was two octaves. The central SF of the moving stimuli was 0.125 or 0.5 cpd, and the RMS contrast was 5%. "Flickering" samples were produced by filtering 32% RMS contrast 1D white noise images. Their central SFs varied from 0.0625 to 4 cpd in octave increments, and samples with higher central SF had higher RMS contrast. Thirty-two randomly interleaved stimulus conditions comprised a single block of trials.

Procedures

Experimental paradigms were controlled by two personal computers (PCs), which communicated via Ethernet (TCP/IP protocol). The first PC utilized Real-Time EXperimentation software (Hays, Richmond, & Optican, 1982), which provided the overall control of the experimental protocol, acquisition, display, and storage of the eye-movement data. The other PC utilized the Psychophysics Toolbox extensions of Matlab (Brainard, 1997; Pelli, 1997) and generated the visual stimuli.

At the start of each trial a stationary stimulus (randomly selected from a lookup table) and a fixation target (diameter = 0.25°) appeared together at the screen center. After the subject's eye had been positioned within 2° of the fixation target and no saccades had been detected (using an eye velocity threshold of 18°/s) for a randomized period of 600 to 1000 ms, the fixation target disappeared and motion commenced. The motion lasted for 200 ms; the screen then turned to uniform gray (luminance = 20.8 cd/m^2), marking the end of the trial. A new stimulus appeared after a 500-ms intertrial interval, signaling a new trial. The subjects were asked to refrain from blinking or shifting fixation except during the intertrial intervals but were given no instructions relating to the motion stimuli. If no saccades were detected for the duration of the trial, then the data were stored; otherwise, the trial was aborted and repeated within the same block. Data collection usually occurred over several sessions until

each condition had been repeated an adequate number of times to permit good resolution of the responses (through averaging; the exact number of trials per condition is indicated in the legends of all figures that present experimental data).

Data analysis

The calibration procedure provided eye-position data that were fitted with second-order polynomials and later used to linearize the horizontal eye-position data recorded during the experiment. Eye-position signals were then smoothed with an acausal sixth-order Butterworth filter (3 dB at 30 Hz), and mean temporal profiles were computed for each stimulus condition. Trials with microsaccadic intrusions (that had failed to reach the eye velocity cut-off of 18°/s used during the experiment) were deleted. To minimize the impact of directional asymmetries and boost the signal-to-noise ratio, the mean horizontal eye position with each leftward motion stimulus was subtracted from the mean horizontal eye position with the corresponding rightward motion stimulus (the mean horizontal eye position). Mean eye velocity was estimated from differences between mean horizontal eye-position samples 10 ms apart (central difference method) and evaluated every 1 ms. Response latency was estimated by determining the time after stimulus onset when the mean eye velocity first exceeded 0.1°/s. The initial OFRs to a given stimulus were quantified by measuring the changes in the mean horizontal eye-position signals (OFR amplitude) over the initial open-loop period—that is, over the period up to twice the minimum response latency. This window always commenced at the same time after the stimulus onset (stimulus-locked measures) and, for a given subject, was the same in all experiments reported in this paper (66–132, 66–132, and 74–148 ms for subjects AGB, BMS, and EJF, respectively).

All error bars in the figures are one standard error of the mean (actually, they were smaller than a symbol size in the vast majority of the OFR amplitude cases and therefore not visible on the graphs). For panel D in Figure A1, a bootstrapping procedure was used to construct 95% confidence intervals.

Results

Figure 1C compares mean eye velocity profiles in response to white noise (dotted black trace) with those in response to white noise stimuli after notch filtering (noted by grayscale coding of velocity traces). These profiles were obtained from subject BMS for stimuli moving at 88°/s. Removing low-SF components produced response reductions: The light-gray traces fall below the response to unfiltered stimuli (dotted trace). Conversely, removing high-SF components led to OFR enhancement, as all dark solid traces lie above the dotted one. The strongest attenuation and enhancement were observed for notch filters centered at approximately 0.18 and 1 cpd, respectively. These features are quantified in Figure 2A through C, which shows how the OFR amplitudes change as a function of the central SF of the notch filter in three subjects. The OFR amplitude change for each speed of motion is calculated relative to the response to unfiltered noise moving at the same speed (the thin horizontal dotted lines indicate no change in response). For all three test speeds, we observe attenuation of the OFR amplitude when low SFs are removed. As speed increases, the attenuation becomes greater and the effect is maximal at lower SF. In two subjects (AGB and EJF), for the lowest speed used (22°/s; blue square symbols), removing high-SF components had no effect. In all other cases the OFR is enhanced by removing high-SF components, and again as speed increases the maximal effect moves to lower SF. Band-pass filtered noise was tested at 44°/s only, and the OFR amplitudes to such stimuli are shown by filled black circles in Figure 2D through F. For comparison, the responses to notch filtered noise moving at the same speed (open black circles in Figure 2A through C) are replotted here as well (open gray circles), showing that the central SFs of a filtered stimulus producing the strongest OFR and a notch producing the strongest OFR attenuation are very similar. On the other hand, in the high SF range where notch filters produced enhancement, the band-pass filtered stimuli produce weak responses. At the highest SFs, responses to band-pass filtered stimuli were often in the reverse direction. This outcome is most likely a manifestation of spatial aliasing: Vertical arrows below the horizontal axes of Figure 2D through F indicate the SF at which aliasing produces the optimal TF in the reverse direction. Note that the magnitude of these reversed responses is much smaller than the magnitude of the enhancement produced by notch filtering, so it seems that spatial aliasing is not responsible for the bulk of this enhancement.

A straightforward qualitative interpretation of all these findings is that high and low SFs play very different roles. It seems like low SFs provide most of the "drive" for the OFR, in a way that straightforwardly reflects the effect of these SF components presented alone. Higher SF components seem to give rise to a form of inhibition that is not seen when these components are presented alone. In order to test these ideas quantitatively, we now explore these two components in more detail.

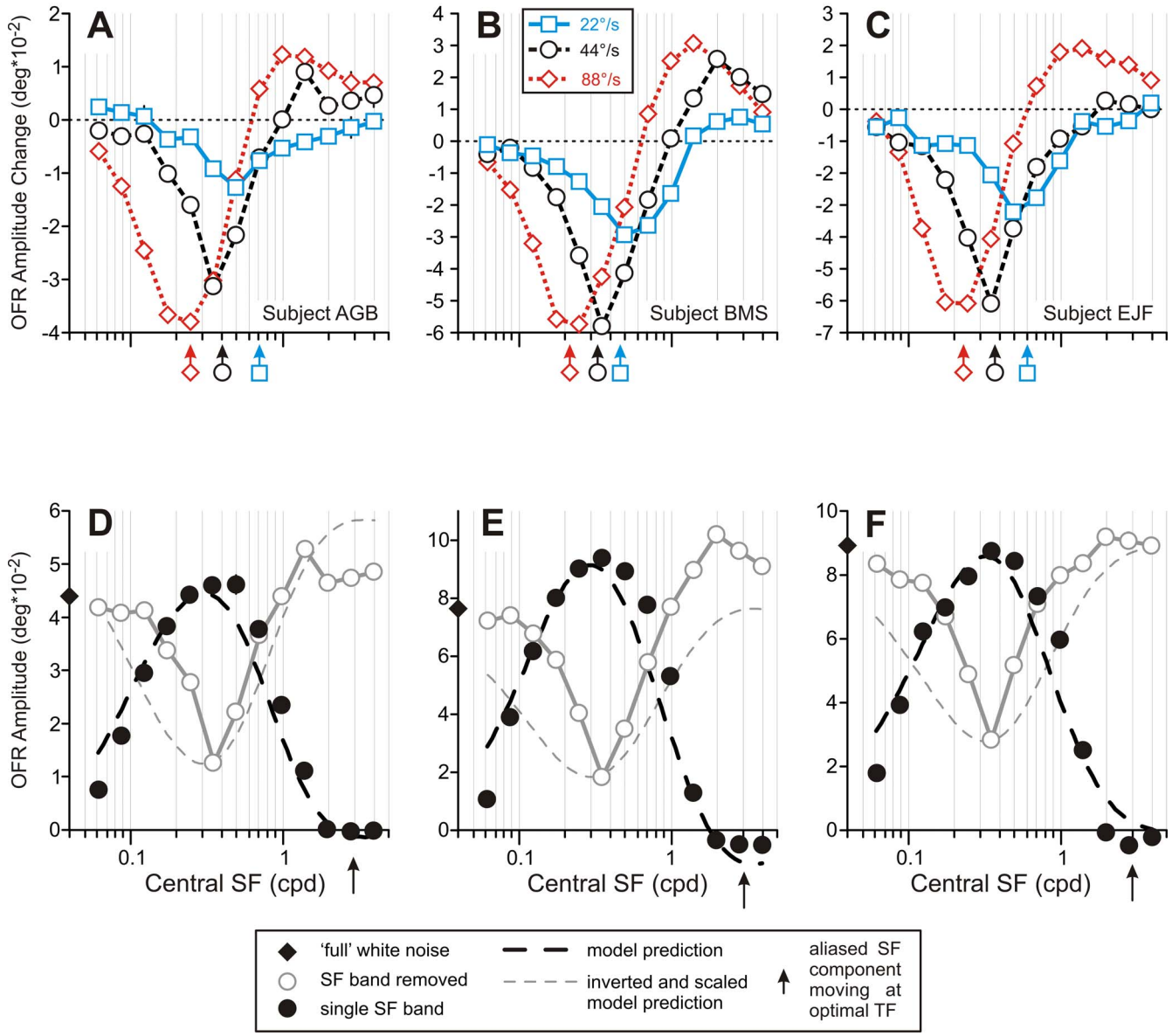

Figure 2. Experiment 1. (A–C) Dependence of mean OFR amplitude on central SF of the notch filter for three speeds of motion: blue squares = 22°/s; black circles = 44°/s; red diamonds = 88°/s. Data are displayed as a change in the OFR amplitude relative to the same-speed unfiltered white noise responses (noted as thin horizontal dotted lines at the ordinate axis zero level). (D–F) Dependence of mean OFR amplitude on central SF of filtered noise stimuli. Filled black circles and diamond = data; black dashed lines = Equation 3 predictions. Additional symbol conventions are shown in the rectangular insert below panels D through F. See main text for details. Subject AGB = panels A and D (172–236 trials/condition); subject BMS = panels B and E (106–161 trials/condition); subject EJF = panels C and F (118–178 trials/condition).

OFR drive

The properties of the OFR have been extensively studied using luminance-modulated sinusoidal gratings, and the OFR magnitude is known to depend on SF (e.g., Sheliga et al., 2005), TF (Gellman et al., 1990), and contrast (e.g., Sheliga et al., 2005). In order to test whether the drive inferred from Figure 2 might be explained by a linear summation of spatiotemporal components, it is necessary to estimate responses to many combinations of SF and TF. This is a great deal simpler if the responses are a separable function of SF and TF, so we first examine this question in Experiment 1B. Such spatiotemporal separability is observed in the majority of simple cells in the striate cortex as well as in a considerable number of complex cells (Priebe et al., 2006) and neurons in areas MT and MST (Miura, Inaba, Aoki, & Kawano, 2014a; Priebe et al., 2003).

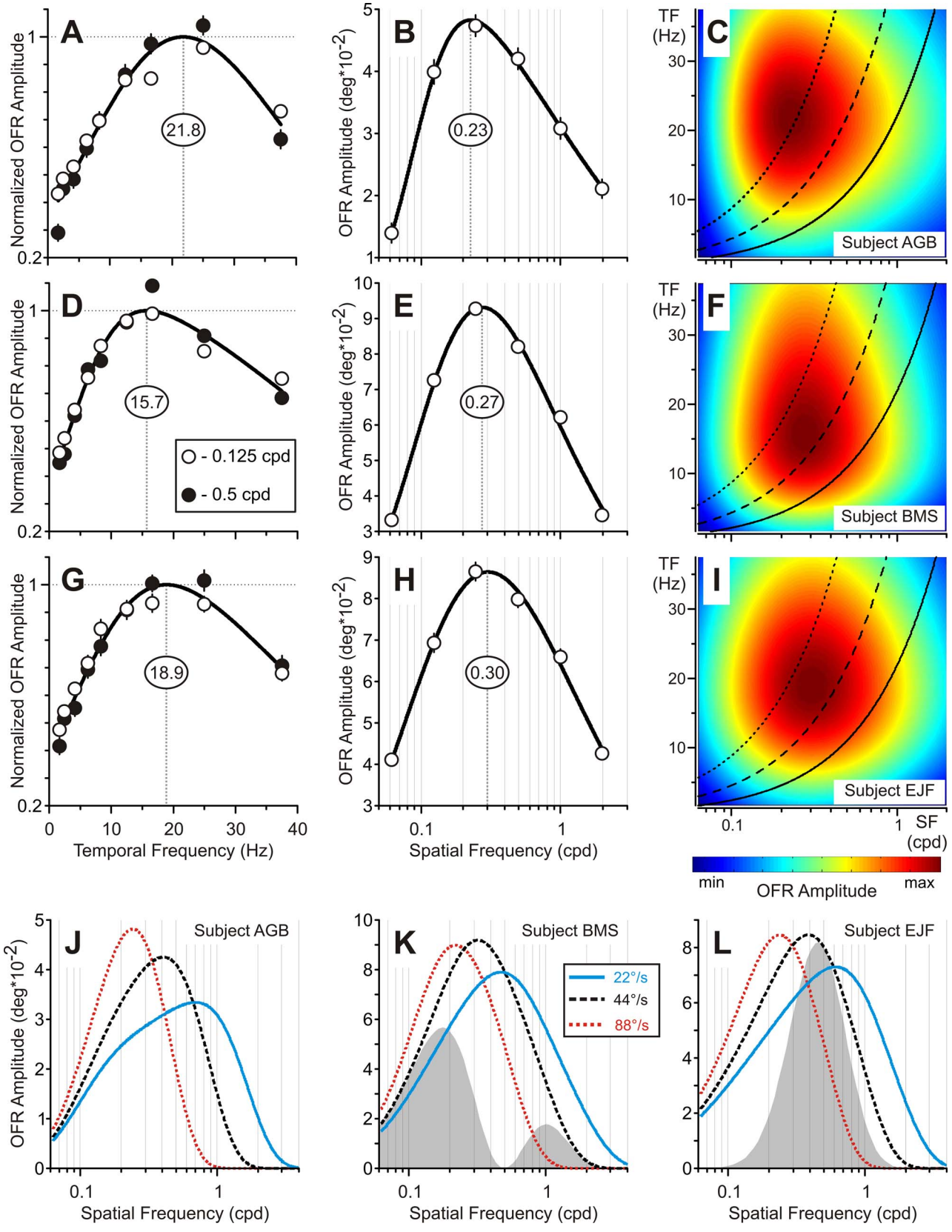

Figure 3. Experiment 1B: Normalized TF tuning curves for sinusoidal gratings of different SFs for subjects AGB (panel A; 99–124 trials/condition), BMS (panel D; 127–135 trials/condition), and EJF (panel G; 68–80 trials/condition). Open circles = 0.125 cpd; filled circles = 0.5 cpd; vertical dotted line = optimal TF. Experiment 1C: SF tuning curves for filtered noise stimuli moving at near-optimal TF for →

←

subjects AGB (panel B; 123–133 trials/condition), BMS (panel E; 193–199 trials/condition), and EJF (panel H; 99–109 trials/condition). Vertical dotted line = optimal central SF. The OFR spatiotemporal surfaces are derived from the TF and SF tuning data for subjects AGB (panel C), BMS (panel F), and EJF (panel I). Black lines on the surface = spatiotemporal frequency profiles (SF-TF profiles) for 22°/s (solid lines), 44°/s (dashed lines), and 88°/s (dotted lines). SF-TF profiles as a function of SF for subjects AGB (panel J), BMS (panel K), and EJF (panel L): solid blue lines = 22°/s; dashed black lines = 44°/s; dotted red lines = 88°/s. Gray shaded areas = examples of the summation areas for notch filtered (panel K) and band-pass filtered (panel L) noise stimuli. See main text for details.

Panels A, D, and G of Figure 3 show the normalized TF tuning curves obtained in Experiment 1B for sinusoidal gratings of different SFs from subjects AGB, BMS, and EJF, respectively.[1] Though SFs of sine wave gratings differed fourfold, their TF tuning was very similar—compare the open and filled circles data—indicating that the responses are well described by a separable function of SF and TF. The TF dependence was well described by a skewed Gaussian, and the strongest OFR amplitudes were recorded at 21.8, 15.7, and 18.9 cycles/s for subjects AGB, BMS, and EJF, respectively. See Supplementary Table S1 for the full list of fit parameters. These results replicate the findings of Gellman et al. (1990), who used continuous rather than apparent motion stimuli.

Given this separable response, we can combine the TF tuning with a measure of SF tuning at a single TF to estimate the full spatiotemporal response surface. In order to do this, we again used band-pass filtered noise, and then set the speed such that the central SF was always moving with the optimal TF. We chose to measure SF tuning with filtered noise to limit the extent to which we depend on an assumption of linear summation over contrast. Since neurons in the striate cortex have bandwidths that increase with SF on the linear scale, white noise stimuli have higher contrast within the passband as SF increases. The RMS contrast of the stimuli used in Experiment 1C followed this pattern so that the results can be interpreted without requiring a strong assumption of a linear contrast response. Panels B, E, and H of Figure 3 show the SF tuning curves obtained in this way. The SF dependence is well described by a skewed Gaussian function of log frequency, and the strongest OFR amplitudes were recorded with central SFs of 0.23, 0.27, and 0.30 cpd for subjects AGB, BMS, and EJF, respectively. Supplementary Table S2 provides the full list of fit parameters. Finally, the full spatiotemporal response surface was constructed by taking the product of the TF and SF tuning data fits under the assumption of spatiotemporal separability and is shown in panels C, F, and I of Figure 3 for subjects AGB, BMS, and EJF, respectively.

The black lines in panels C, F, and I of Figure 3 show the spatiotemporal frequency profiles of white noise stimuli moving at 22°/s (solid lines), 44°/s (dashed lines), and 88°/s (dotted lines)—that is, the motion speeds used in Experiment 1.[2] These same profiles are shown in Figure 3J through L as a function of SF. For each motion speed, a slightly different SF produces the strongest drive (given by the peaks of the curves). The locations of these peaks (for each subject) are marked by color- and symbol-coded arrows located below Figure 2A through C. These lie very close to the minimum in the response to a notch filtered stimulus, confirming that this notch reflects the simple withdrawal of drive. Figure 2A through C also shows that larger speeds produce deeper troughs; this is also reflected in the drive estimates from Figure 3J through L. Taken together, these findings suggest that a relatively simple summation of spatiotemporal components accounts for the OFR drive produced by white noise stimuli. Note, however, that a purely linear summation is unlikely to provide a good account: The width of the trough produced by notch filtering is considerably smaller than the width of the function describing excitatory responses to band-pass stimuli.[3] This may reflect the winner-take-all behavior governing the combination of components of different contrasts (e.g., Sheliga et al., 2006), a possibility we return to below.

OFR inhibition

OFRs to white noise stimuli lacking high-SF components were stronger than those to unfiltered noise (Figure 2A through C). This suggests that high-SF components exert inhibitory influences on mechanisms driving an OFR. Importantly, this inhibition appears to reflect a nonlinear interaction: The band-pass stimuli at these SFs do not produce substantial negative responses. One possible candidate for this inhibition might be a consequence of using apparent motion stimuli, which contain a series of fixed-amplitude spatial displacements applied for a number of successive video frames. If this step size is large relative to the receptive field size of a direction-selective neuron, then the receptive field stimulation is equivalent to stimulation with a different noise sample on each video frame, which we call *flicker*. Although flicker does not produce a consistent motion signal, it may activate the channels in a way that can interfere with other motion signals. In order to test this possibility explicitly, in Experiment 1D we constructed

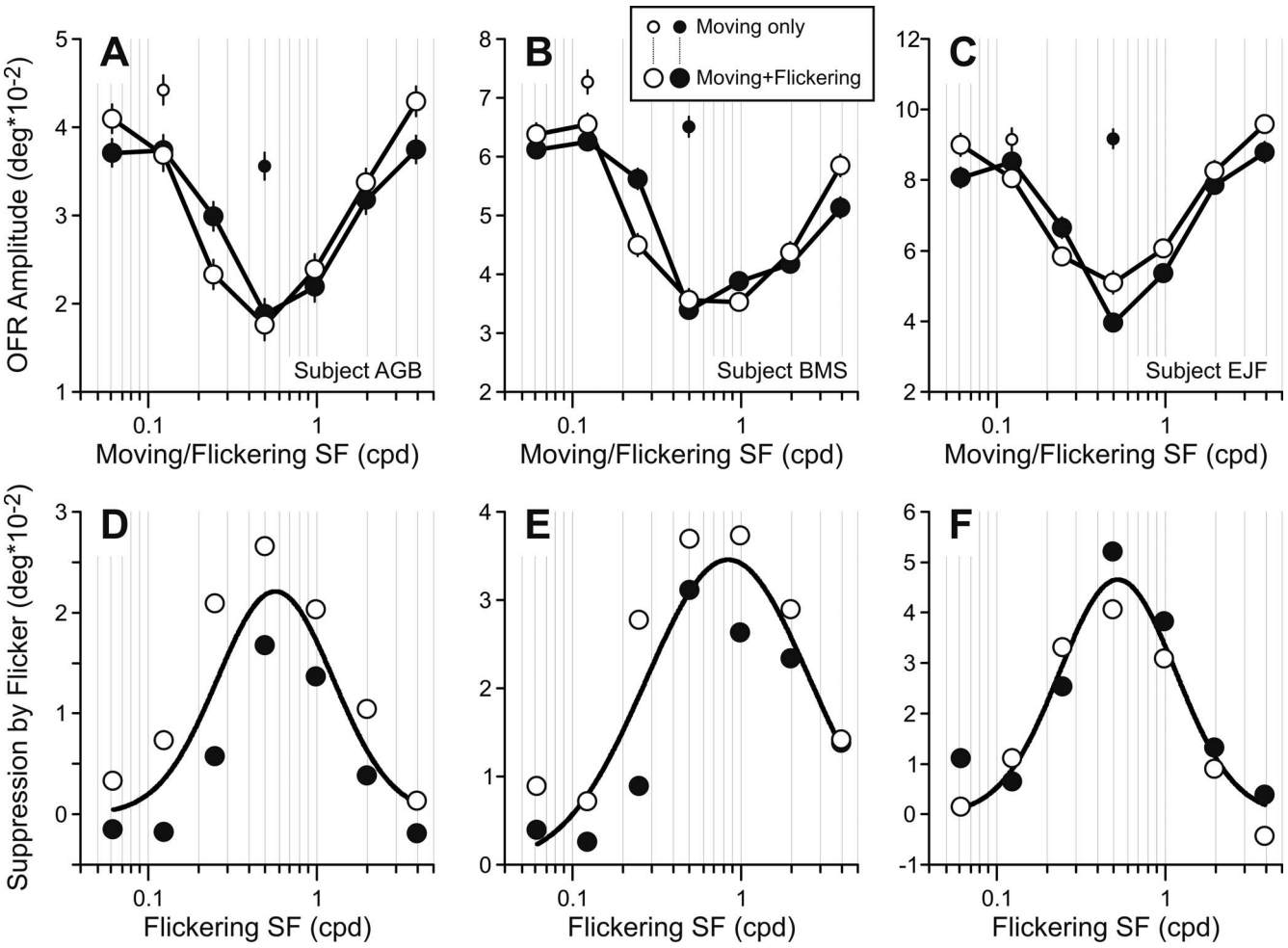

Figure 4. Experiment 1D. (A–C) Dependence of mean OFR amplitude on flicker SF. Small circles = moving band-pass filtered noise only; large circles = combined moving and flickering filtered noise. Central SF of moving filtered noise: open circles = 0.125 cpd; filled circles = 0.5 cpd. (D–F) The OFR amplitude suppression (absolute drop) due to flicker. Solid lines = Gaussian fits for pooled data. Subject AGB = panels A and D (152–167 trials/condition); subject BMS = panels B and E (174–187 trials/condition); subject EJF = panels C and F (76–99 trials/condition).

stimuli by summing one moving pattern (band-pass filtered noise at a near-optimal TF) and one flickering pattern (band-pass filtered noise, where we selected a different noise pattern on each video frame).[4] Figure 4A through C shows that the addition of flickering stimuli (large circles) greatly reduced the OFRs' responses to moving stimuli (small circles show responses to a moving stimulus in isolation). The effect depended on the SF of the flickering stimulus: Flickering stimuli with central SFs in the range of 0.5 to 1 cpd were the most detrimental. The effect was largely unaffected by changes in the SF of the moving stimulus. Figure 4D through F plots change in OFR amplitude as a function of flickering central SF. Pooled data were described satisfactorily by a single semilog Gaussian fit ($r^2 = 0.760$, 0.804, and 0.925 for subjects AGB, BMS, and EJF, respectively). See Supplementary Table S3 for the full list of fit parameters.

A simple model

The experiments above identify two factors that, when combined, may explain the responses observed in Figure 2. First, there is an excitatory drive, possibly involving a winner-take-all competition between components. We model this by a power law summation (Britten & Heuer, 1999):

$$\mathrm{OFR}_{DRIVE} = \left(\sum_{i=1}^{k} D_i^n \right)^{\frac{1}{n}}, \quad (1)$$

where D_i are observed responses to individual Fourier components, and the exponent (n) determines the strength of the winner-take-all rule ($n = 1$ corresponding to purely linear summation). At each point the value of D_i is derived by multiplying the observed response to a given SF (using the fits in Figure 3J through L) by the contrast of that frequency in the

filtered image.[5] The shaded gray area in Figure 3K illustrates the values of D_i in the notch filtered stimulus moving at 44°/s. The shaded gray area in Figure 3L illustrates the values of D_i in the band-pass filtered stimulus moving at 44°/s.

Second, there is an inhibitory effect of flicker, which arises when the discrete steps in the image movement are large relative to the spatial filters that detect motion at a given spatial scale. In order to model this, we first have to define the extent of this flicker for each spatial filter given a stimulus. One simple way to do this is to examine the phase shifts corresponding to the image step size. For filters of finite SF bandwidth, these phase shifts will not be identical for different components within the passband. We quantify the scatter in these phase shifts simply by computing the circular variance (CV) of the Fourier components of two filtered images following a single apparent motion step. That is, the image step is represented in the Fourier domain with a set of vectors whose length represents the Fourier amplitude and whose angle represents the phase shift for that component. We then compute the CV of these vectors. CV is a common measure of dispersion in circular statistics (e.g., Berens, 2009), where a value of 0 indicates no dispersion (all vectors of the sample point in the same direction) and a value of 1 indicates a uniform circular distribution (the vectors are spread out evenly around the circle). For white noise moving with a fixed speed, the CV will also depend on filter bandwidth.[6] We chose to fix this at the mean of physiologically observed values using a Gaussian in log frequency with an FWHM of 1.4 octaves (De Valois et al., 1982). Note that in filters of fixed octave bandwidth, the CV is larger in higher SF filters, even when comparing two stimuli with the same TF applied to the central SF of the filter.

Given only an estimate of filter bandwidth, CV defines the extent of flicker produced at each SF in any stimulus, shown for our 1D noise stimuli in Figure 5A. Higher SF channels experience more flicker than low-SF channels, and flicker increases with stimulus speed (a given stimulus speed in Figure 5 is coded by both line type and its color). Note that these functions are determined only by the stimulus and the presumed filter bandwidth. We then combine this with our empirical measures of inhibition by flicker (Figure 4, where our stimulus ensures that CV = 1 at all frequencies in the flickering component) to estimate the impact of flicker in any given stimulus:

$$\mathrm{OFR}_{FLICKER} = \sum_{i=1}^{k} [CV_i * F_i], \quad (2)$$

where F_i are the empirically observed values of Fourier components, obtained from the fits in Figure 4D through F, deconvolved with the filter that was applied in generating the flickering stimulus. Values of F_i are shown as green dotted-dashed lines in Figure 5B through D.[7] A point-by-point multiplication of CV_i and F_i—shown in Figure 5B through D by blue solid, black dashed, and dotted red lines for three speeds in each of the three subjects—results in the dependencies showing how much loss in the OFR amplitude should be expected due to flicker of various SF components. The shaded areas in Figure 5C and D illustrate the values of $CV_i{*}F_i$ in a 1-cpd notch filtered stimulus moving at 88°/s (red shading; Figure 5C) and a 1-cpd band-pass filtered stimulus moving at 44°/s (gray shading; Figure 5D).

Finally we describe the OFR responses with the difference of Equations 1 and 2 after applying a scaling coefficient to the flicker component:

$$OFR = \mathrm{OFR}_{DRIVE} - K_F{*}\ \mathrm{OFR}_{FLICKER}. \quad (3)$$

Applying this model to each subject's data requires estimating just two free parameters: n and K_F. The model fits to OFRs in unfiltered and notch filtered stimuli are shown in Figure 6, where it produces an excellent account of the data in Experiment 1: $r^2 =$ 0.927, 0.929, and 0.932 for subjects AGB, BMS, and EJF, respectively. The same model also describes the responses to moving white noise passed through different spatial band-pass filters (dashed black lines in Figure 2D through F). We also use Equation 3 to describe the data in Experiments 2 and 3 below (Figures 7 and 8). In generating these fits, we found the single pair of parameters for each subject that provided the best description of all these data (Figures 6, 7, and 8). Table 1 provides the best-fit values of free parameters shared by all three experiments.

Discussion

Experiment 1 demonstrated contrasting roles played by different Fourier components in moving white noise stimuli generating OFRs. Lower SF Fourier components drive the OFRs, while higher SFs produce inhibition. The drive from low SFs is largely determined by the response to the strongest single component in a separable function describing combinations of SF and TF. Thus, the SF at which notch filtering produces the most attenuated response matches the peak in the tuning to the spatiotemporal components in that stimulus (peaks from Figure 3J through L plotted with arrows in Figure 2A through C). However, the width of the troughs produced by notch filtering did not match the width of this spatiotemporal tuning: The latter were considerably wider ($p < 0.004$; paired t test using data for 44°/s, $n = 3$). We quantified this by comparing their FWHMs; the median ratio was 0.66.[8] This indicates that the contribution of the most

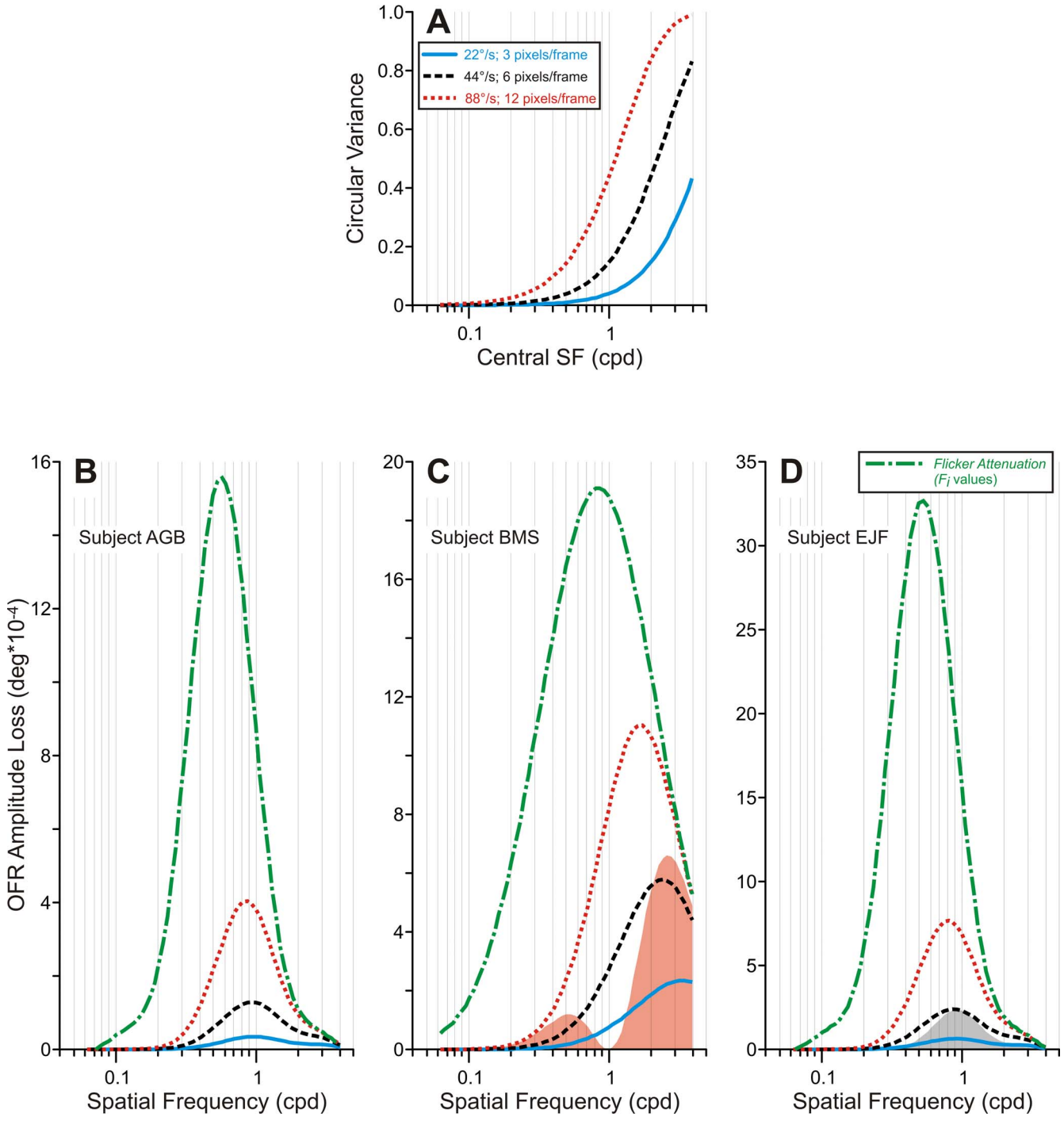

Figure 5. Modeling the OFR suppression by flicker. Three speeds of stimulus motion: 22°/s (blue solid lines), 44°/s (black dashed lines), and 88°/s (red dotted lines). (A) The dependencies between the central SF of a motion detector and extent of flicker produced by apparent motion (quantified by CV, CV_i). (B–D) A loss in the OFR amplitude expected due to flicker of various Fourier components for subjects AGB (panel B), BMS (panel C), and EJF (panel D). The F_i values are the empirically observed values of Fourier components, obtained from the fits in Figure 4D through F, deconvolved with the filter that was applied in generating the flickering stimulus; they are shown as green dotted-dashed lines. A point-by-point multiplication of CV_i and F_i (shown by blue solid, black dashed, and dotted red lines) results in the dependencies showing how much loss in the OFR amplitude should be expected due to flicker of various SF components. Shaded areas = examples of the summation areas for 1-cpd–centered notch filtered (red area, panel C) and 1-cpd–centered band-pass filtered (gray area, panel D) noise stimuli. See main text for more details.

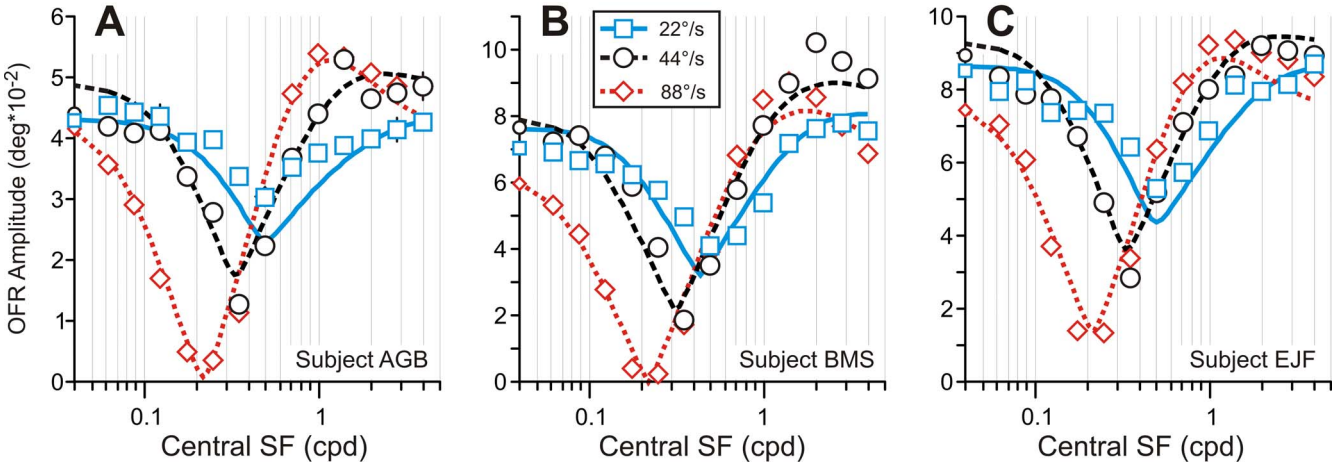

Figure 6. Experiment 1. Dependences of mean OFR amplitude on central SF of the notch filter for three speeds of motion: data (open blue squares = 22°/s; open black circles = 44°/s; open red diamonds = 88°/s) and Equation 3 (solid blue lines = 22°/s; dashed black lines = 44°/s; dotted red lines = 88°/s) fits. The OFRs to unfiltered noise stimuli are shown on the ordinate axis and are appropriately color and symbol coded. Subject AGB = panel A; subject BMS = panel B; subject EJF = panel C.

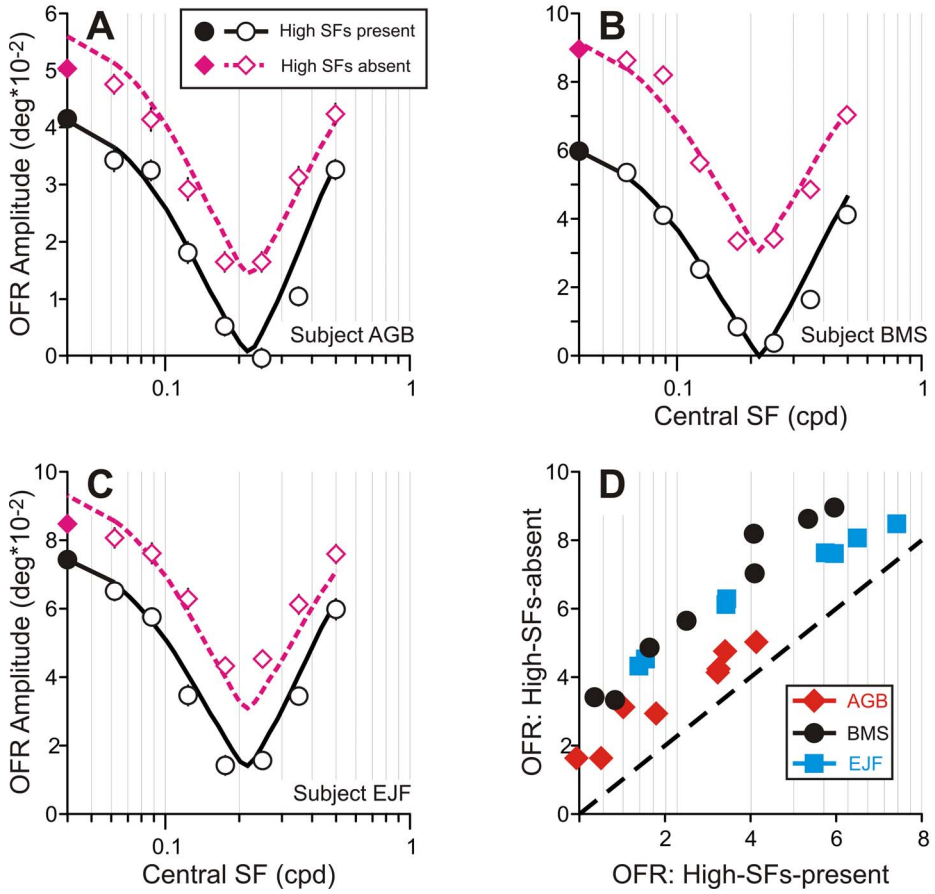

Figure 7. Experiment 2. Dependences of mean OFR amplitude on central SF of the notch filter for high SFs absent (open pink diamonds, dashed lines) and high SFs present (open black circles, solid lines) sets of stimuli. Data (symbols) and Equation 3 (lines) fits. The OFRs to noise stimuli that were not notch filtered are shown on the ordinate axis (filled symbols) and are appropriately color and symbol coded. Subject AGB = panel A (136–156 trials/condition); subject BMS = panel B (114–123 trials/condition); subject EJF = panel C (88–116 trials/condition).

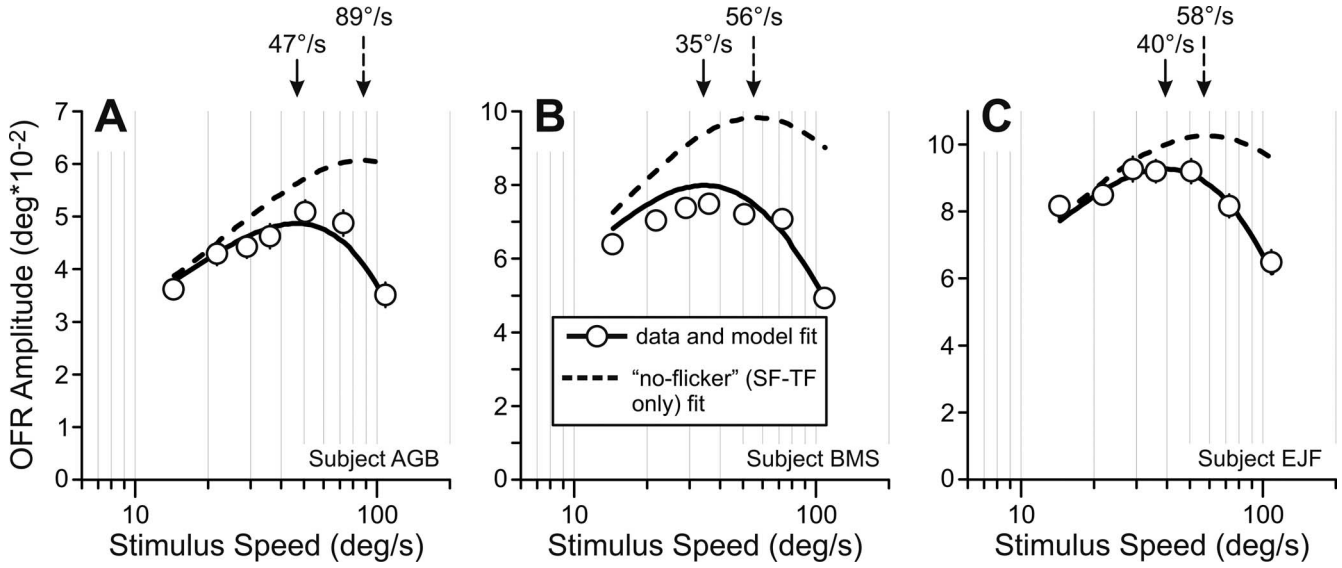

Figure 8. Experiment 3. The OFR speed tuning. Solid lines = Equation 3 fits; dashed lines = SF-TF only fits (i.e., assuming no flicker contribution). Numbers and arrows above the panels indicate stimulus speeds at which the fits place the maximal OFR: solid arrows = Equation 3 fits; dashed arrows = no-flicker fits. Subject AGB = panel A (95–103 trials/condition); subject BMS = panel B (192–199 trials/condition); subject EJF = panel C (66–85 trials/condition).

dominant Fourier component toward the OFR magnitude was disproportionally high, resembling a winner-take-all behavior. These two major findings of Experiment 1 parallel neurophysiological observations made in area MT of primates: When tested with random dot stimuli moving at different speeds, most MT neurons showed the same peak but narrower tuning than predicted by the spatiotemporal response surfaces—that is, by applying the linear model to the responses to sine wave gratings across a wide range of spatial and temporal frequencies (Priebe et al., 2003). On the other hand, the vast majority of V1 neurons did not show this narrowing of speed tuning with random dots (Priebe et al., 2006), suggesting that a winner-take-all–like mechanism acts after V1. Quantitatively, deviations from the linear model predictions were slightly stronger in our study than those found in the MT (median FWHM ratio of 0.66 vs. 0.87), but it is quite possible that further sharpening of the winner-take-all–like effects occurs in visual areas beyond the MT (e.g., in area MST, which was shown to be crucial for OFR generation; Takemura, Murata, Kawano, & Miles, 2007).

			Goodness of fit (r^2)		
Subject	n	K_F	Experiment 1	Experiment 2	Experiment 3
AGB	8.9	2.69	0.927	0.943	0.864
BMS	30.9	1.23	0.929	0.969	0.692
EJF	12.2	1.84	0.932	0.938	0.939

Table 1. Experiments 1, 2 and 3: best-fit parameters for Equation 3. n = power term, K_F = flicker scaling coefficient.

Because apparent motion stimuli contain discrete image displacements ("steps"), this can disrupt motion signals in higher SF channels. Although this is a well-recognized property of temporally sampled displays, we identify here a novel nonlinearity, which means that components that lie outside the window of visibility can profoundly influence responses to moving stimuli. We propose that this reflects a consequence of temporal decorrelation in spatially finite filters: Large steps effectively produce flicker in small filters. We tested this possibility by generating true decorrelation over a set of SFs in the stimulus and confirmed that this reduces the amplitude of OFRs even though they produce no response when presented alone. A simple model based on this observation produced an excellent account of the responses to notch filtered stimuli across the entire range.

The model required only two free parameters: the exponent, n, and the scaling coefficient for flicker, K_F. The value of parameter n, though quite variable, is higher than 8 for all subjects, commensurate with a winner-take-all behavior in the summation of OFR drive. The value of parameter K_F varied twofold in our three subjects. One possible interpretation of this variability is that there is an effective difference in channel bandwidth between subjects. In an alternative formulation of the model we can remove the free parameter K_F and instead fit the channel bandwidth used when calculating CV. In two subjects this produced reasonable values for the bandwidth (2.1 in BMS, 2.6 in EJF), but did not do so in the third.

However, in the absence of other evidence for intersubject differences in bandwidth, we feel it would be premature to commit to this form of description in the model.

A peculiar feature of our model, however, is that it postulates that the outputs of the mechanisms driving and suppressing the OFRs are subtracted from one another (Equation 3). This is puzzling because many previous studies of interactions between Fourier components report nonlinear interactions (Britten & Heuer, 1999; Carandini & Heeger, 1994; Carandini, Heeger, & Movshon, 1997; Heuer & Britten, 2002; Kumbhani et al., 2008; Liu & Sperling, 2006; Miura et al., 2014b; Sheliga et al., 2006; Simoncelli & Heeger, 1998). Indeed, it seems improbable that subtraction can be a good description in all conditions. If, for example, the strength of the driving stimulus were reduced to produce a drive that was smaller than the flicker term in Equation 3, it is hard to believe that adding the flickering stimulus would reverse the direction of the OFRs. However, we explored several other ways of incorporating the effects of flicker (e.g., multiplicative scaling of the OFR drive, modification of a contrast normalization scheme), and none provided good descriptions of the data. We therefore did an additional experiment to test the idea that this interaction was really subtractive.

Simoncini, Perrinet, Montagnini, Mamassian, and Masson (2012) varied the bandwidth of band-pass random-texture stimuli and did not report any response reduction caused by flicker. This probably reflects two differences between our stimuli and theirs: (a) The bandwidth of their stimuli was much narrower (FWHM was always less than two octaves; range = 0.2–1.9), and (b) the frequencies were not very high. As a result, our estimate of flicker strength would be low in their stimuli. For the purposes of the study by Simoncini et al. (2012), it was advantageous not to elicit flicker. But their results are entirely compatible with ours.

Experiment 2

In order to test the idea of a subtractive interaction between drive and flicker, we sought to evaluate the effect of adding flicker to a range of stimuli that produced different response magnitudes. We wished to do this while keeping the stimuli as close as possible to those used in Experiment 1. A simple way to do this is to compare the effect of notch filtering the stimulus with or without the high-SF components. Therefore, in Experiment 2 we compared OFRs to two sets of white noise stimuli moving at a certain speed. In all stimuli composing the first set, higher SFs (which at this particular speed of motion would develop considerable flicker) were removed, whereas higher SFs were preserved in stimuli of the second set. We then applied notch filters at various central SFs, exactly as in Experiment 1. Where high SFs have been removed, the notch filters simply remove driving frequencies and hence reduce responses. According to our proposed subtractive mechanism, adding high SFs back into these stimuli should produce a reduction in OFR of the same magnitude regardless of the initial response strength. If the flickering components acted through a normalization mechanism [e.g., $OFR_{NORM} \sim D/(1 + F)$], then the effect of flicker (in absolute units) should decline with the driving response.

Materials and method

Only methods and procedures that were different from those used in Experiment 1 are described.

Visual stimuli

For each subject, vertical 1D 32% RMS contrast white noise stimuli (1 pixel wide) were randomly generated. The speed of motion was approximately 88°/s (12 pixels/frame) for all three subjects (we used a high speed to maximize the effect of flickering components). Two sets of stimuli were created. One was identical to those used in Experiment 1 ("high SF present"). The second was low-pass filtered, removing all components with SF greater than 0.71 cpd ("high SF absent"). Both types of stimuli were then passed through notch filters as in Experiment 1. Each set also included a stimulus without the notch filtering, resulting in 32 randomly interleaved stimuli in a single block of trials.

Data analysis

Experiments 1 and 2 had several conditions in common, one of which was the unfiltered white noise stimulus moving at 88°/s. To allow quantitative comparisons between two experiments, we normalized all OFR measures obtained in Experiment 2 with respect to the OFR measure recorded in this common condition.

Results and discussion

Figure 7A through C plots OFR amplitudes as a function of the central SF of the notch filters for both sets of noise stimuli: high SFs present (black circles) and high SFs absent (pink diamonds). Equation 3—our subtraction mechanism—provided excellent fits for data of all subjects: $r^2 = 0.943$, 0.969, and 0.938 for

subjects AGB, BMS, and EJF, respectively (both sets of data were fitted by a single fitting procedure). That these changes are better described by subtraction than multiplication is most clearly seen in Figure 7D, which plots responses in each high SF present condition against responses in the equivalent high SF absent condition. These data lie parallel to a line of slope 1 but have a fixed offset relative to the identify line.

Experiment 3

In Experiment 3 we measured responses to unfiltered white noise moving at a range of different speeds. We had two objectives: (a) to see whether our model provided a good account of the observed speed tuning in these stimuli and (b) to use the resulting fits to estimate the impact of this flicker response on the observed speed tuning.

Materials and method

Only methods and procedures that were different from those used in Experiment 1 are described.

Visual stimuli

Unfiltered vertical 1D white noise stimuli moved horizontally at a range of different speeds: approximately 15°, 22°, 29°, 36°, 51°, 73°, and 109°/s (i.e., 2, 3, 4, 5, 7, 10, and 15 pixels/frame, respectively). A single block of trials had 14 randomly interleaved stimuli: seven speeds and two directions of motion (left vs. right).

Data analysis

Experiments 1 and 3 had one condition in common: unfiltered white noise moving at 22°/s. To allow quantitative comparisons between two experiments, we normalized all OFR measures obtained in Experiment 3 with respect to the OFR measure recorded in this common condition.

Results

Figure 8 plots the OFR amplitudes as a function of speed, and Equation 3 provided good fits in all subjects (solid lines); $r^2 = 0.864$, 0.692, and 0.939 for subjects AGB, BMS, and EJF, respectively. The fit for subject BMS did slightly overestimate the OFR amplitudes at low speeds, but this feature was already evident in Experiment 1. (Note a discrepancy between the data point—blue square at the ordinate axis—and the fit for the 22°/s stimulus in Figure 6B.) The best-fit speed tuning curves peaked at 47°, 35°, and 40°/s for subjects AGB, BMS, and EJF, respectively. For each subject these peak values were much lower than those predicted simply from the drive that reflects summation of spatiotemporal components (89°, 56°, and 58°/s for subjects AGB, BMS, and EJF, respectively; dashed lines).

Discussion

CRT monitors are commonly used to deliver visual stimulation in laboratory settings.[9] This inevitably means that we must use apparent motion stimuli (discrete image steps) to study behavioral and neural mechanisms of visual motion processing. It is often assumed that, provided a high-enough frame rate, this distinction is unimportant because the window of visibility imposed by the display encompasses the range of the human visual system. Here we report a novel nonlinearity, which has the result that components with temporal frequencies beyond the Nyquist limit of the display can systematically alter the pattern of results. In particular, Experiment 3 documents an important example: Estimates of speed tuning can be significantly distorted by the flicker that apparent motion produces in high-SF channels despite the fact that we use a relatively high frame rate (150 Hz). At lower stimulus refresh rates, these differences should be even greater.[10] How could one minimize these detrimental effects of the apparent motion? Increasing the refresh rate of the CRT monitor is an obvious solution. But this approach has its limitations because the vast majority of (still) available CRT monitors cannot support refresh rates higher than 150 Hz, even at the lowest spatial resolutions. Another approach would be to diminish the power of the high-SF components of a broadband stimulus. The 1D white noise stimuli of our study assigned a "black" or "white" value to each successive pixel along the horizontal axis. If we were to randomly assign the same value ("black" or "white") to quartets of neighboring pixels along the axis of motion—4-pixel-wide noise—it would lead to dramatic changes in such stimulus' 1D FFT, removing power from high frequencies (compare the upper panel of Figure 1B and Figure 1D). It would lead to a much weaker flicker and, therefore, much less interference from higher SF components toward lower SF components. In the case of 2D noise, increasing the size of single elements (e.g., checks or dots) would have the same effect. Finally, one might consider using pink noise stimuli in place of white noise stimuli. An additional potentially beneficial advantage of pink noise stimuli in many applications is

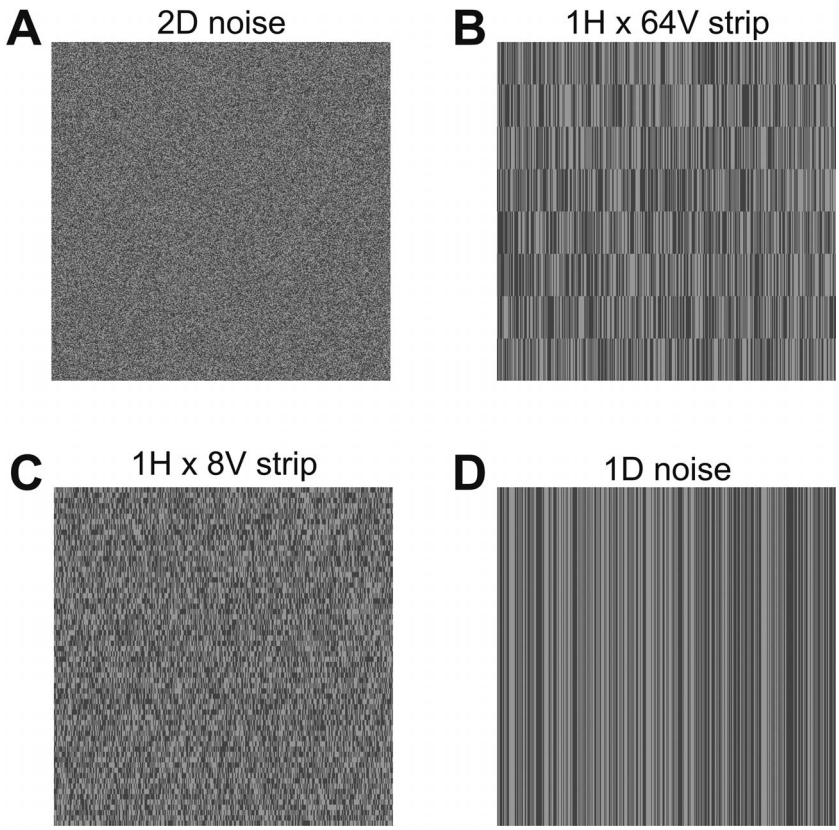

Figure 9. Experiment 4. Examples of vertical 2D (A–C) and 1D (D) white noise stimuli: scaled versions of a 25° × 25° 1-pixel-wide patterns. Stimuli comprised one or more abutting horizontal strips of vertical 1D noise. Examples include strip heights of approximately 0.05° (A), approximately 0.4° (B), approximately 3° (C), and approximately 25° (D).

that they follow natural images' statistics much more closely.

Experiment 4

All of the previous experiments used 1D stimuli. Even though the rules governing the interactions between Fourier components may be the same in 2D stimuli, other nonlinearities may change the effective contribution of different frequencies. Thus, recently we studied the OFRs to horizontal motion of vertical sine wave gratings whose total size was fixed but who comprised a variable number of abutting horizontal strips where alternate strips were in counterphase (Sheliga, Quaia, FitzGibbon, & Cumming, 2013). We found that for any one SF, the OFRs varied with strip height, and the optimum strip height was proportional to stimulus wavelength; that is, as the strip height increased, lower SFs played a stronger role. Changes in strip height might therefore alter the balance between drive and flicker. In noise patterns, varying strip height spans the space between 1D noise and 2D noise. When the strip height equals the pixel width, the stimulus is 2D noise; when the strip height equals the stimulus height, it becomes 1D noise. Intermediate stimuli can also be constructed by using abutting 1D noise strips of variable height (see Figure 9). Experiment 4 therefore set out to evaluate the interaction between strip height and speed tuning to determine whether the interactions we report in 1D stimuli also influence responses in 2D stimuli.

Materials and method

Only those methods and procedures that were different from those used in Experiment 1 are described.

Visual stimuli

White noise stimuli comprised one or more abutting horizontal strips of vertical 1D noise. Vertical 1D noise was 1 pixel wide; each successive pixel along the horizontal axis was randomly assigned a "black" or "white" value. Strip height varied from approximately 0.05° to approximately 25° (i.e., from 1 to 512 pixels) in octave increments. Stimuli occupied a rectangular

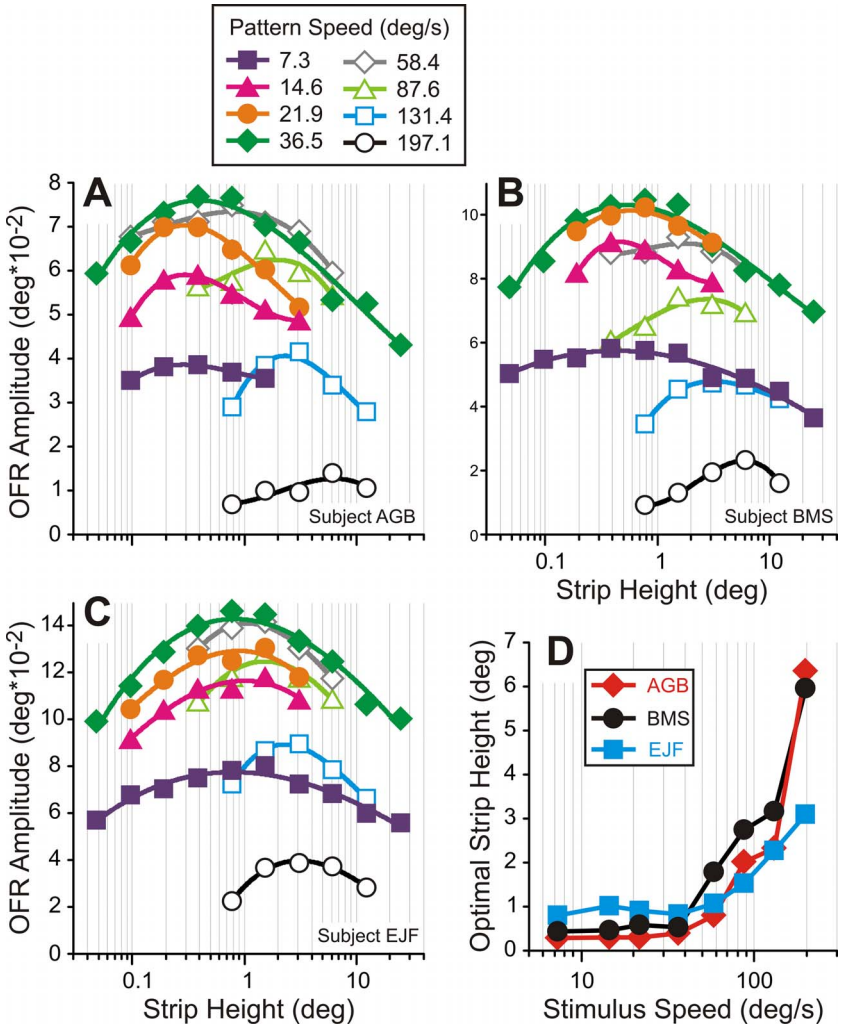

Figure 10. Experiment 4. (A–C) Dependences of mean OFR amplitude on the strip height of white noise stimulus. Different motion speeds are color and symbol coded (see the insert). (D) The dependence between the optimal strip height and the speed of motion. Subject AGB = panel A (108–174 trials/condition); subject BMS = panel B (142–199 trials/condition); subject EJF = panel C (106–160 trials/condition).

aperture (approximately 25° × 25°) centered at fixation. Figure 9 shows examples of stimuli; see the figure legend for details. Several speeds of motion were tested: approximately 7°, 15°, 22°, 36°, 58°, 88°, 131°, and 197°/s (i.e., 1, 2, 3, 5, 8, 12, 18, and 27 pixels/frame, respectively). Not every combination of strip height and speed of motion was implemented. For every subject we collected pilot data to select conditions that would span the optimal speed at each strip height. All stimuli had 32% RMS contrast. Different strip height, speed, and direction of motion conditions were randomly interleaved in a single block (98–104 trials/block).

Results

Figure 10A through C plots the relationship between strip height and OFR amplitude for eight different speeds of noise motion for three subjects.[11] There are three major features. First, the OFR amplitudes exhibited a great deal of variation with speed, simply reflecting speed tuning similar to that shown in Figure 8. Second, the dependences on strip height for any given speed were bell shaped, with a clear optimum at each speed. This indicates that changes in strip height are changing the effectiveness of different Fourier components. Third, the optimal strip height—that resulted in the largest OFR—shifted toward larger values for higher speeds of motion. This latter result is summarized in Figure 10D, which plots the optimal strip height as a function of speed. Results are quite similar for three subjects and are characterized by a rather flat relationship for speeds up to approximately 36°/s followed by a rise in the optimal strip height for higher speeds of motion. Since the OFR is driven by lower SF components at higher speeds (Experiment 1),

this probably indicates that larger strip heights preferentially activate low-SF channels. In the Appendix we test whether the OFRs observed in Experiment 4 could be quantitatively accounted for by our subtractive model (Equation 3; one-subject data analysis).

General discussion and concluding remarks

Measurements of speed tuning curves to motion of white noise stimuli constitute an integral part of many quantitative studies of motion sensitivity in perception as well as in the description of neurons in different visual cortices. In apparent motion displays, high spatial frequencies often move with TFs that exceed the temporal Nyquist limit of a given display. Although these components produce no response alone, we demonstrate an important new nonlinearity: These flickering components suppress responses produced by lower SF components. The effect of this nonlinearity in single-unit studies may depend on the stage of visual processing in which this nonlinearity is implemented. The earliest stages of visual processing might not be affected as much as the later ones, where more nonlinear interactions between SF components are observed.

We show that for any given speed of motion, the spatiotemporal properties of the OFRs are the sole determinant of which Fourier components will contribute the most toward the OFR drive. We also demonstrate that the contribution of the strongest components toward the OFR magnitude is disproportionally high—a winner-take-all behavior. Despite the fact that white noise stimuli are broadband, the SF channel that is most activated almost completely dominates the response. Neurophysiological evidence suggests that this pattern of activity is already present in the extrastriate MT (Priebe et al., 2003) but not in the primary visual cortex (Priebe et al., 2006).

Given the evidence for winner-take-all behavior, the OFR inhibition caused by flicker of higher SF components is surprising. The higher SF components make relatively weak contributions and therefore should effectively be eliminated by the winner-take-all stage. It is also puzzling that the interaction was best described by subtraction because the drive signal is directional, whereas the flicker is not. To our knowledge, there are no neurophysiological demonstrations of such subtractive interactions, though one article's modeling of psychophysical judgments required subtractive interaction to account for direction judgements at short latency (Serrano-Pedraza, Goddard, & Derrington, 2007). Some aspects of our data—for example, that the optimal SF for flicker (~1 cpd) is substantially higher than that for drive (compare Figure 4A through C with Figure 3B, E, and H; also see the Appendix)—already indicate that different mechanisms subserve drive and flicker. It is therefore possible that the mechanism subserving flicker is not affected by the winner-take-all process. This suggests that in broadband stimuli such as those used here, the interaction between components is more complex than the simple winner-take-all behavior evident when just two grating components are used (Sheliga, Fitzgibbon, & Miles, 2008; Sheliga et al., 2006). A more complex form of winner-take-all in broadband stimuli might even explain behavior that looks subtractive. The flickering stimulus contains motion energy in both directions. Consequently, if some winner-take-all–like competition were to diminish the effect only of energy in the same direction as the moving stimulus, the remaining signal would appear subtractive.

It seems likely that flicker and drive are combined differently when used for speed perception rather than OFRs. Jogan and Stocker (2015) found that human perceptual judgements of motion of a mix of several band-pass stimuli whose central SFs differed eightfold were best described by a Bayesian model of optimal integration, quite unlike the subtraction we observed. Simoncini et al. (2012) described opposite dynamics of changes in the OFR versus perceptual sensitivity as the bandwidth of their random-texture stimuli was increased. Finally, Wexler, Glennerster, Cavanagh, Ito, and Seno (2013) reported that a flicker delivered after a period of slow motion led to a perception of high-speed jump.

In conclusion, we have demonstrated that a novel nonlinearity produces an artifact when using moving white noise stimuli with discrete temporal sampling of the display. This gives rise to artificial signals that would not be present during continuous motion, as in real-world settings, especially at high speeds of motion. The impact of this nonlinearity can be limited by carefully controlling the SF content of noise stimuli.

Keywords: visual motion, visual flicker, speed tuning, broadband visual stimuli

Acknowledgments

This research was supported by the Intramural Research Program of the National Eye Institute at the National Institutes of Health.

Commercial relationships: none.
Corresponding author: Boris M. Sheliga.
Email: bms@nei.nih.gov.
Address: Laboratory of Sensorimotor Research, National Institutes of Health, Bethesda, MD, USA.

Footnotes

[1] In absolute terms, the OFRs to 0.5 cpd sine wave stimuli were smaller than those to 0.125 cpd stimuli in all three subjects. The normalized plots hide this aspect of the data: For each SF, the data were fit by a skewed Gaussian function and then normalized in respect to the maximal value of the fit. Fits shown in Figure 3A, D, and G take into account the normalized amplitudes of OFRs to both tested sine wave stimuli.

[2] These are not straight lines because the abscissa (but not the ordinate) is on a log scale.

[3] Figure 2D through F provides an example. For a stimulus moving at 44°/s, the SF tuning curves for band-pass stimuli (black dashed lines) were considerably wider than the troughs produced by notch filtering (gray symbols and solid lines). To facilitate the comparison, the SF tuning curves for band-pass stimuli are superimposed (gray dashed lines) after inverting and scaling.

[4] Note that because these samples were produced by filtering 32% RMS contrast 1D white noise images using a Gaussian envelope on a log scale, samples with higher central SFs had higher RMS contrasts.

[5] Because D_i was derived from the fitted functions in Figure 3J through L, we used a fine sampling of frequency components, from 2^{-6} to 2^4 cpd, spaced evenly in log frequency (every 0.1 $\log_2$ cpd unit; i.e., k = 101). We used logarithmic spacing because our fitted functions were on a logarithmic scale.

[6] Phase shifts of various Fourier components were weighted according to their power (Gaussian envelope squared).

[7] F_i was evaluated from the fits for the same spatial frequencies as D_i.

[8] Range = 0.52 to 0.86 (see Supplementary Table S4).

[9] The phenomena that we describe apply equally to liquid crystal and digital light processing displays, but these are less commonly used to study motion in the laboratory.

[10] In a pilot experiment we measured the OFR speed tuning curves utilizing a 75-Hz stimulus refresh rate rather than the 150-Hz rate used in all experiments of this manuscript. This results in larger values of CV and, hence, stronger flicker, which should reduce responses at high speeds. The data showed peaks at 33°, 25°, and 26°/s for subjects AGB, BMS, and EJF, respectively. These values are lower than those obtained in Experiment 3, as expected from our model.

[11] The data were fit by the third-order polynomial functions (see Supplementary Table S5 for the full list of fit parameters).

[12] Phases of gratings in different strips were randomized.

[13] In absolute terms, the OFRs to gratings of 6.2° strip height were smaller than those to gratings comprising strips of the other two heights. For each strip height condition, the data were fitted by a skewed Gaussian and then normalized in respect to the maximal value of the fit. The fit shown in panel A ($r^2 = 0.963$) takes into account normalized amplitudes of OFRs to stimuli of all three strip height conditions. See Supplementary Table S6 for the full list of fit parameters.

[14] Best-fit values of free parameters were $n = 118.3$ and $K_F = 1.07$.

References

Berens, P. (2009). CircStat: A Matlab toolbox for circular statistics. *Journal of Statistical Software, 31*(10), 1–21.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10,* 433–436.

Britten, K. H., & Heuer, H. W. (1999). Spatial summation in the receptive fields of MT neurons. *The Journal of Neuroscience, 19,* 5074–5084.

Carandini, M., & Heeger, D. J. (1994). Summation and division by neurons in primate visual cortex. *Science, 264,* 1333–1336.

Carandini, M., Heeger, D. J., & Movshon, J. A. (1997). Linearity and normalization in simple cells of the macaque primary visual cortex. *The Journal of Neuroscience, 17,* 8621–8644.

Collewijn, H., Van Der Mark, F., & Jansen, T. C. (1975). Precise recording of human eye movements. *Vision Research, 15,* 447–450.

De Valois, R. L., Albrecht, D. G., & Thorell, L. G. (1982). Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research, 22,* 545–559.

De Valois, R. L., & De Valois, K. K. (1988). *Spatial vision*. Oxford, United Kingdom: Oxford University Press.

Gellman, R. S., Carl, J. R., & Miles, F. A. (1990). Short latency ocular-following responses in man. *Visual Neuroscience, 5,* 107–122.

Hays, A. V., Richmond, B. J., & Optican, L. M. (1982). A UNIX-based multiple process system for real-time data acquisition and control. *WESCON Conference Proceedings, 2,* 1–10.

Heuer, H. W., & Britten, K. H. (2002). Contrast dependence of response normalization in area MT of the rhesus macaque. *Journal of Neurophysiology, 88,* 3398–3408.

Jogan, M., & Stocker, A. A. (2015). Signal integration

in human visual speed perception. *The Journal of Neuroscience, 35,* 9381–9390.

Kumbhani, R. D., Saber, G. T., Majaj, N. J., Tailby, C., & Movshon, J. A. (2008). Contrast affects pattern direction selectivity in macaque MT neurons. *Program No. 460.26*, 2008 Neuroscience Meeting Planner. Washington, DC: Society for Neuroscience, 2008. Online.

Liu, D., & Sperling, G. (2006). Motion strength is not what is summed in the vector summation computation of plaid motion. *Journal of Vision, 6*(6): 1046, doi:1010.1167/1046. [Abstract]

Masson, G. S., & Perrinet, L. U. (2012). The behavioral receptive field underlying motion integration for primate tracking eye movements. *Neuroscience & Biobehavioral Reviews, 36,* 1–25.

Miles, F. A. (1998). The neural processing of 3-D visual information: Evidence from eye movements. *The European Journal of Neuroscience, 10,* 811–822.

Miles, F. A., Kawano, K., & Optican, L. M. (1986). Short-latency ocular following responses of monkey. I. Dependence on temporospatial properties of visual input. *Journal of Neurophysiology, 56,* 1321–1354.

Miles, F. A., & Sheliga, B. M. (2010). Motion detection for reflexive tracking. In U. Ilg & G. S. Masson (Eds.), *Dynamics of visual motion processing: Neuronal, behavioral and computational approaches* (pp. 141–160). New York, NY: Springer-Verlag.

Miura, K., Inaba, N., Aoki, Y., & Kawano, K. (2014a). Difference in visual motion representation between cortical areas MT and MST during ocular following responses. *The Journal of Neuroscience, 34,* 2160–2168.

Miura, K., Inaba, N., Aoki, Y., & Kawano, K. (2014b). Responses of MT/MST neurons elicited by dual-grating stimulus: Differences between areas MT and MST. *Program No. 726.709*. 2014 Neuroscience Meeting Planner. Washington, DC: Society for Neuroscience, 2014. Online.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10,* 437–442.

Priebe, N. J., Cassanello, C. R., & Lisberger, S. G. (2003). The neural representation of speed in macaque area MT/V5. *The Journal of Neuroscience, 23,* 5650–5661.

Priebe, N. J., Lisberger, S. G., & Movshon, J. A. (2006). Tuning for spatiotemporal frequency and speed in directionally selective neurons of macaque striate cortex. *The Journal of Neuroscience, 26,* 2941–2950.

Robinson, D. A. (1963). A method of measuring eye movement using a scleral search coil in a magnetic field. *IEEE Transactions in Biomedical Engineering, 10,* 137–145.

Serrano-Pedraza, I., Goddard, P., & Derrington, A. M. (2007). Evidence for reciprocal antagonism between motion sensors tuned to coarse and fine features. *Journal of Vision, 7*(12):8, 1–14, doi:10.1167/7.12.8. [PubMed] [Article]

Sheliga, B. M., Chen, K. J., FitzGibbon, E. J., & Miles, F. A. (2005). Initial ocular following in humans: A response to first-order motion energy. *Vision Research, 45,* 3307–3321.

Sheliga, B. M., Fitzgibbon, E. J., & Miles, F. A. (2008). Spatial summation properties of the human ocular following response (OFR): Evidence for non-linearities due to local and global inhibitory interactions. *Vision Research, 48,* 1758–1776.

Sheliga, B. M., Kodaka, Y., FitzGibbon, E. J., & Miles, F. A. (2006). Human ocular following initiated by competing image motions: Evidence for a winner-take-all mechanism. *Vision Research, 46,* 2041–2060.

Sheliga, B. M., Quaia, C., FitzGibbon, E. J., & Cumming, B. G. (2013). Retinal visual processing constrains human ocular following response. *Vision Research, 93,* 29–42.

Sheliga, B. M., Quaia, C., FitzGibbon, E. J., & Cumming, B. G. (2014a). Speed tuning of human ocular following responses (OFRs) depends on orientation bandwidth in noise stimuli. *Journal of Vision, 14*(10): 477, doi:10.1167/14.10.477. [Abstract]

Sheliga, B. M., Quaia, C., FitzGibbon, E. J., & Cumming, B. G. (2014b). Two mechanisms are required to explain ocular-following responses (OFRs) in humans to white noise stimuli. *Program No. 726.11*, 2014 Neuroscience Meeting Planner. Washington, DC: Society for Neuroscience, 2014. Online.

Simoncelli, E. P., & Heeger, D. J. (1998). A model of neuronal responses in visual area MT. *Vision Research, 38,* 743–761.

Simoncini, C., Perrinet, L. U., Montagnini, A., Mamassian, P., & Masson, G. S. (2012). More is not always better: Adaptive gain control explains dissociation between perception and action. *Nature Neuroscience, 15,* 1596–1603.

Takemura, A., Murata, Y., Kawano, K., & Miles, F. A. (2007). Deficits in short-latency tracking eye movements after chemical lesions in monkey cortical areas MT and MST. *The Journal of Neuroscience, 27,* 529–541.

Wexler, M., Glennerster, A., Cavanagh, P., Ito, H., & Seno, T. (2013). Default perception of high-speed motion. *Proceedings of the National Academy of Sciences, USA, 110,* 7080–7085.

Yang, D. S., FitzGibbon, E. J., & Miles, F. A. (2003). Short-latency disparity-vergence eye movements in humans: Sensitivity to simulated orthogonal tropias. *Vision Research, 43,* 431–443.

Appendix

In one subject (BMS), we attempted to see whether the observed OFRs could be quantitatively accounted for by our subtractive model (Equation 3). To obtain the data required for such an analysis, we ran Experiments 1B through 1D using stimuli that all had the same total size (approximately 25° × 25°) but comprised abutting strips of variable height. Panel A of Figure A1 shows the normalized TF tuning (like in Experiment 1B) obtained for 0.5-cpd sinusoidal gratings arranged in 0.4°- (pink open diamonds), 1.6°- (blue open squares), or 6.2°- (green filled diamonds) high strips.[12] Although gratings' strip heights differed 16-fold, their tuning was similar: The response is a separable function of TF and strip height. The strongest OFRs were recorded at 18.6 cycles/s.[13] Panel B of Figure A1 shows the SF tuning curves (like in Experiment 1C) for filtered noise stimuli whose strip heights ranged from 0.1° to 25° in two-octave increments (see the Figure A1 insert for symbol definition). All dependences were fit by skewed Gaussian functions (median $r^2 = 0.998$, range = 0.984–1.000; see Supplementary Table S7 for the full list of fit parameters) and revealed that, as the strip height increased, the peaks of the dependencies shifted toward lower central SFs (similar to findings of Sheliga et al., 2013). Panel C of Figure A1 shows results of an experiment similar to Experiment 1D but run using filtered noise stimuli whose strip heights ranged from

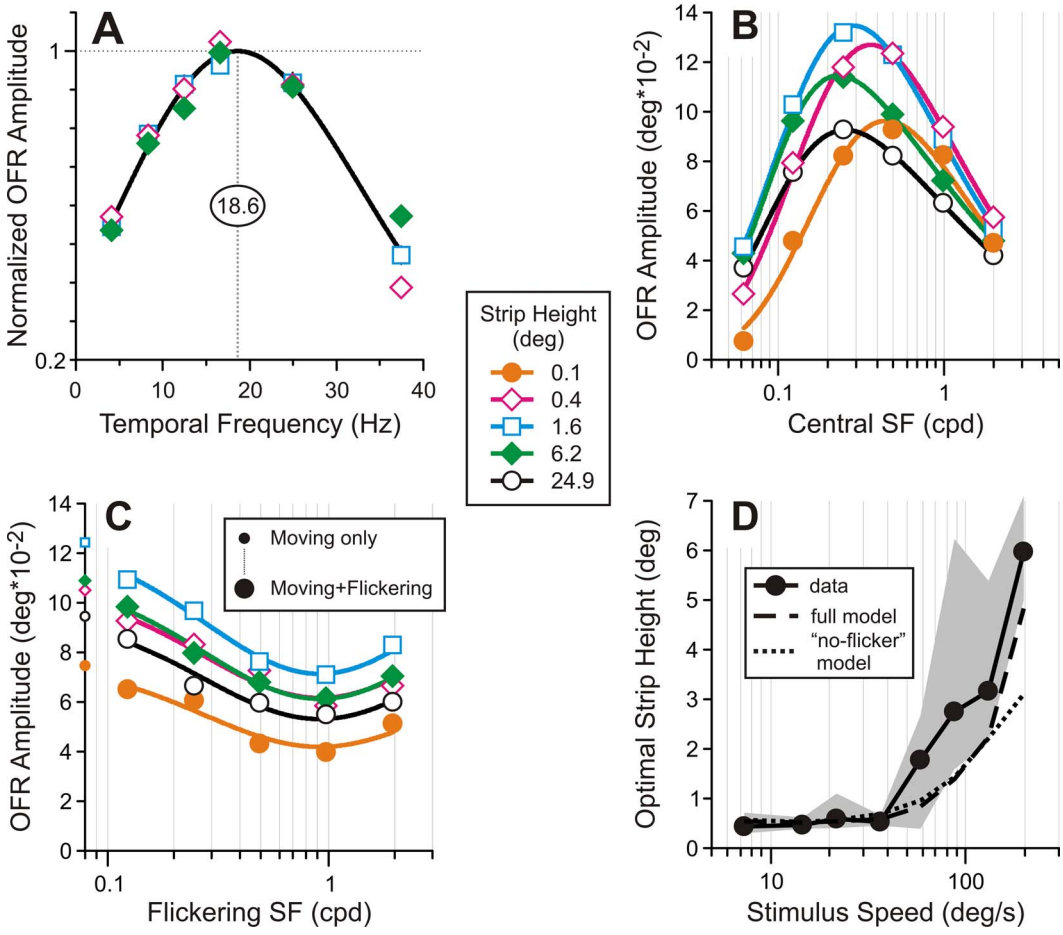

Figure A1. Experiment 4, subject BMS. (A) OFR TF tuning; 81 to 85 trials/condition. (B) OFR SF tuning; 93 to 100 trials/condition. (C) The OFR amplitude dependence on flicker SF; 93 to 100 trials/condition. Different strip heights are color and symbol coded (see the insert). (D) The dependence between the optimal strip height and the speed of motion. Filled circles and solid line = data; gray shaded area = 95% confidence interval; dashed line = Equation 3 prediction; dotted line = no-flicker prediction.

0.1° to 25° in two-octave increments. Recall that in this experiment the pattern on the screen was a sum of two filtered noise stimuli: The first moved with the near-optimal TF, whereas the second was a new randomly chosen filtered noise sample substituted each frame (i.e., flicker). Panel C shows that an addition of flickering stimuli (larger symbols) strongly reduced the OFRs compared with moving stimuli presented in isolation (smaller symbols; 0.25 cpd central SF, shown on the ordinate axis). This effect depended on SF but did not depend on strip height. The data were well fit by semilog Gaussians (median $r^2 = 0.943$, range = 0.908–0.987). See Supplementary Table S8 for the full list of the best-fit Gaussian parameters. Constraining the Gaussian peak SF and sigma parameters to be the same for stimuli of different strip height had minor impact; it led to a 1.1% (range = 0.1%–4.4%) drop in r^2 values. The best-fitting Gaussian amplitudes, on the other hand, varied considerably with strip height (range = 0.033°–0.053°) and, in fact, were linearly related to the OFR amplitudes to moving stimuli of different strip heights when those were presented in isolation ($r^2 =$ 0.978; not shown). The changes in the amplitude of attenuation may reflect changes in the effective SF components that drive flicker, parallel to the changes in drive. Using the data from panels A through C of Figure A1, it was possible to use Equation 3 to predict the data for subject BMS in Experiment 4 data ($r^2 =$ 0.957; the entire data set was evaluated by a single fitting procedure).[14] The model dependence between the optimal strip height and speed of motion is shown in panel D of Figure A1 using a black dashed line. This figure also replots subject BMS's experimentally measured curve from Figure 10D (filled circles and solid lines) along with its 95% confidence interval (gray shaded area). The model fit provides a good account of the data, although it underestimates the steepness of the rise at the largest speeds. We also show the result of using the model without incorporating the effect of flicker (black dotted line), which produces similar results over most of the range. This suggests that there is little change in the balance of flicker and drive with strip height so that the effects we demonstrated with 1D noise (Experiments 1 through 3) seem to have similar effects across this range of 2D stimuli.

The data in Figure A1 provide two indications that the processing of these flicker signals goes through different channels than the motion signals that drive the response. First, the optimal SF for flicker (~1 cpd) is substantially higher than that for drive (this is also seen comparing Figure 4 with Figure 3). Second, while changes in strip height cause changes in the preferred SF for driving responses, the preferred SF for flicker does not change with strip height. We do not have adequate data to make well-founded suggestions about what differences in mechanisms give rise to these differences in behavior.