The spread of attention across features of a surface

Zachary Raymond Ernst,[1] Geoffrey M. Boynton,[1] and Mehrdad Jazayeri[2]

[1]*Department of Psychology, University of Washington, Seattle, Washington; and* [2]*Howard Hughes Medical Institute, National Primate Research Center, and Department of Physiology and Biophysics, University of Washington, Seattle, Washington*

Ernst ZR, Boynton GM, Jazayeri M. The spread of attention across features of a surface. *J Neurophysiol* 110: 2426–2439, 2013. First published July 24, 2013; doi:10.1152/jn.00828.2012.—Contrasting theories of visual attention have emphasized selection by spatial location, individual features, and whole objects. We used functional magnetic resonance imaging to ask whether and how attention to one feature of an object spreads to other features of the same object. Subjects viewed two spatially superimposed surfaces of random dots that were segregated by distinct color-motion conjunctions. The color and direction of motion of each surface changed smoothly and in a cyclical fashion. Subjects were required to track one feature (e.g., color) of one of the two surfaces and detect brief moments when the attended feature diverged from its smooth trajectory. To tease apart the effect of attention to individual features on the hemodynamic response, we used a frequency-tagging scheme. In this scheme, the stimulus features (color and direction of motion) are modulated periodically at distinct frequencies so that the contribution of each feature to the hemodynamics can be inferred from the harmonic response at the corresponding frequency. We found that attention to one feature (e.g., color) of one surface increased the response modulation not only to the attended feature but also to the other feature (e.g., motion) of the same surface. This attentional modulation was evident in multiple visual areas and was present as early as V1. The spread of attention to the behaviorally irrelevant features of a surface suggests that attention may automatically select all features of a single object. Thus object-based attention may be supported by an enhancement of feature-specific sensory signals in the visual cortex.

attention; visual object; fMRI; frequency tagging; transparent motion; color

SELECTIVE ATTENTION IMPROVES information processing for a subset of relevant stimuli, usually at the expense of irrelevant stimuli. Attention can select a region of space (spatial attention), a stimulus feature (feature attention), or a whole object (object attention). What distinguishes object- and feature-based attention is that object-based attention improves processing of all features of a selected object with little or no additional cost. For example, when asked to monitor multiple features simultaneously, subjects are more accurate when the attended features belong to the same object compared with when they are from different objects (Blaser et al. 2000; Duncan 1984; Rodriguez et al. 2002). Despite its importance in behavior, little is known about the mechanisms by which object-based attention influences the representation of individual features in the brain.

A common challenge in studying the mechanisms of object-based attention in humans is that existing tools such as functional magnetic resonance imaging (fMRI) do not have the requisite resolution to tease apart the representation of individual features when they overlap in space and time. Pattern classification techniques have provided a means to circumvent this problem by extracting information from the pattern of hemodynamic blood oxygen level-dependent (BOLD) responses across voxels (Boynton 2005a). For example, in occipital cortex, pattern classification can extract information about orientation (Kamitani and Tong 2005), directions of motion (Kamitani and Tong 2006), and color (Brouwer and Heeger 2009; Kamitani and Tong 2005, 2006). Moreover, this technique has been used to demonstrate how attention to a specific feature can selectively and reliably modulate the pattern of fMRI responses to that feature (Kamitani and Tong 2005, 2006; Serences and Boynton 2007).

Pattern classification methods use sophisticated algorithms to decode information that is not immediately accessible at the level of the spatially averaged BOLD signal. An alternative to this strategy, and one that we have used here, is to design stimuli in ways that would allow information about individual features to be readily encoded by the amplitude of the BOLD signal. To do so, we employed the so-called frequency-tagging technique (Regan 1989), which has been previously used in EEG recordings (Andersen et al. 2008; Muller et al. 2006; Schoenfeld et al. 2007). In this technique, the presentation of each stimulus feature is modulated in time at a specific temporal frequency so that the evoked response associated with that feature can be extracted from the strength of the corresponding harmonic response at that frequency. Accordingly, the response evoked by multiple features can be readily teased apart by tagging each feature with its own unique frequency.

We implemented this strategy in a stimulus that consisted of two superimposed transparent surfaces, each comprising a field of dots with a distinct color-motion conjunction that changed smoothly with time. The direction of motion of one surface changed in a clockwise manner, and the direction of motion of the other surface changed in a counterclockwise manner. Likewise, the color of each surface cycled through our color-space in opposite directions. Frequency tagging was performed by making both the direction of motion and the color of the dots in each surface change periodically, and with distinct frequencies. We then measured the amplitude of the BOLD response at the four frequencies corresponding to the four surface features. We used this experimental setting to determine the effect of feature- and object-based attention on BOLD signals throughout visual cortex. By analyzing the amplitude of the BOLD response at each of the four designated frequencies, we found that attention modulated the response to the attended feature as well as the task-irrelevant feature associated with the same surface. This effect, which was present in multiple visual areas including V1, demonstrates that object-based attention

modulates feature-specific representations across the visual cortex.

METHODS

Participants. Four men and four women ages 20 to 28 yr gave written consent to participate in this study in accord with a protocol approved by the Human Subjects Division at the University of Washington. Subjects all had normal or corrected-to-normal vision, and 6 of them were naive to the purpose of the experiment. Subjects participated in two separate experiments: *1*) an fMRI experiment that consisted of one retinotopic mapping session followed by two 2-h functional scanning sessions, and *2*) a psychophysical experiment that consisted of two 1-h behavioral sessions. Four of the eight subjects participated in both experiments. For both fMRI and psychophysical experiments, subjects completed 1–2 h of training to ensure that they were familiar with the task.

Stimulus. The stimulus used in both the fMRI and behavioral experiments consisted of two superimposed fields of dots. Each dot field consisted of 101 dots per frame (frame rate = 60 Hz) of the same color that moved coherently in a specific direction at a speed of 6 deg/s. The two dot fields had distinct color-motion conjunction and appeared as two surfaces moving transparently across one another (Fig. 1*A* and Supplemental Movie 1; supplemental material for this article is available online at the *Journal of Neurophysiology* website). To remove a potential depth cue, the depth order of each dot (which dots occlude the other dots) was randomized. The two fields were rendered on a black background within an annulus with an inner diameter of 3° and an outer diameter of 16° of visual angle.

At stimulus onset, the dots associated with one of the dot fields (hereafter, *surface 1*) appeared blue and moved leftward within the annulus. At the same time, the dots of the other dot field (hereafter, *surface 2*) appeared red and moved upward. During the presentation of the stimulus, the color and direction of motion of both dot fields changed slowly and cyclically.

Color specifications. The color of dots in each surface and at each time point was determined by a point in the CIE L*a*b* space. Changes in color with time were governed by slow movements of this point along a circular path through the CIE L*a*b* space (Fig. 1, *B* and *C*). The CIE L*a*b* space was chosen for its perceptual uniformity to generate a color sequence that changed in chromaticity at a roughly constant rate. The two surfaces rotated along the same circular path but in opposite directions and with different temporal periods (Fig. 1*B*); *surface 1* went from blue to red to green with a temporal period of 19.20 s (T_{c1}), and *surface 2* went from red to green to blue with a temporal period of 17.14 s (T_{c2}). In the fMRI experiment, in which the stimulus was presented for a total of 8 min, *surfaces 1* and *2* made 19 and 28 full rotations, respectively. The circular path was defined mathematically as follows:

$$L^* = 100$$

$$a_i^*(t) = 42 \times \cos\left(2\pi\frac{t}{T_{ci}} + \phi_i + n\right)$$

$$b_i^*(t) = 42 \times \sin\left(2\pi\frac{t}{T_{ci}} + \phi_i + n\right) \tag{1}$$

where *i* indexes each surface (*i* = 1 or 2), *t* represents elapsed time in seconds, ϕ is a phase parameter that determines the surface's initial hue, and *n* corresponds to the brief dispersions that were added during a color event (see *Color events*). The constant 42 (amplitude of a* and b*) was chosen to keep the colors within the dynamic range of our projector. After specifying the L*a*b* values, we used standard CIE *XYZ* coordinate transformations to compute the corresponding RGB values to drive the calibrated projector.

To ensure that various colors were perceived as isoluminant, we adjusted the scaling of the RGB values using the so-called "minimum
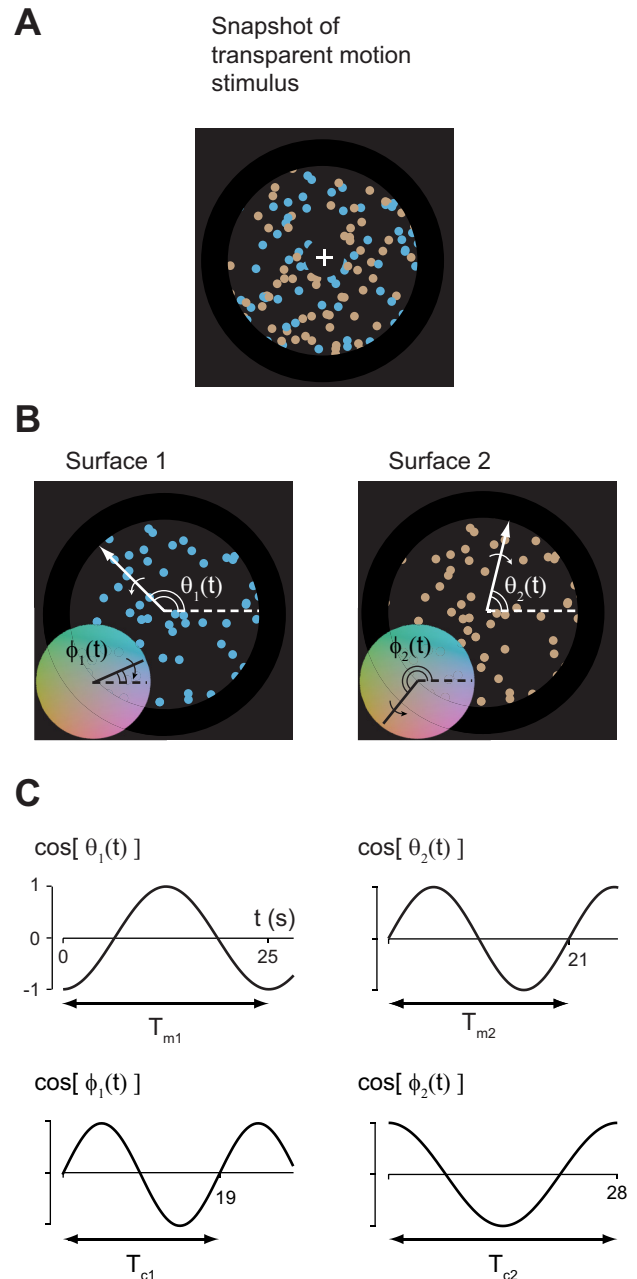

Fig. 1. Description of the stimulus. *A*: a snapshot of the stimulus consisting of 2 superimposed surfaces, each created from a field of dots confined to an annulus centered on a white fixation cross at the center of the screen. *B* shows the 2 surfaces separately. In *surface 1* (*left*), the dots are bluish and are moving up and to the left (large white arrow), and in *surface 2* (*right*), the dots are reddish and moving up and to the right. The color and the direction of motion of the 2 surfaces were modulated continuously and periodically. For *surface 1*, the direction of motion changed in a counterclockwise fashion (small curved white arrow). At a given point in time, *t*, the direction of motion of the dots was specified by $\theta_1(t)$, the angle from the dashed white line. The color of the dots in *surface 1* corresponds to the angle of the black line in the color space (*inset*). The color progressed clockwise through the color space as specified by $\phi_1(t)$, the angle from the dashed black line. For *surface 2*, the direction of motion (large white arrow) progressed clockwise (small curved white arrow) and was specified by $\theta_2(t)$. The color progressed counterclockwise through the color space and was specified by $\phi_2(t)$. *C* shows the first 28 s of the cosine of the 4 angles that specified the color and direction of motion in the stimulus. For *surface 1* (*left*), the temporal periods for the direction of motion and color were 25 s (T_{m1}) and 19 s (T_{c1}), respectively. For *surface 2* (*right*), the temporal periods for the direction of motion and color were 19 s (T_{m2}) and 28 s (T_{c2}), respectively.

motion procedure" (Anstis and Cavanagh 1983). Subjects viewed a stimulus that was made of a chromatic test grating superimposed on a luminance grating. The two gratings were in quadrature phase and were modulated sinusoidally. The test grating was created from two phosphors modulating in antiphase with a temporal period of 333 ms (20 frames) and one phosphor held at a constant intermediate intensity. The luminance grating was made of in-phase space-time modulations of the three phosphors at a Michelson contrast of 0.08. This stimulus is typically perceived as having clockwise or counterclockwise apparent rotational motion unless the two modulating phosphors in the test grating are isoluminant.

We determined this point of isoluminance by a staircase procedure in which subject adjusted the ratio of the two test phosphors so as to null the apparent motion. On each trial, subjects fixated a central spot and used two buttons to reduce or increase the intensity of one of the two modulating phosphors of the chromatic grating (test phosphor) by a fixed amount (step size) to reverse the perceived direction of the apparent motion. After each reversal, the step size was halved and the procedure was repeated until the subject reported that they no longer perceived a clear apparent motion. We quantified this point of isoluminance by the ratio of the intensity of the test phosphor to the other modulating phosphor. To estimate the points of isoluminance across the whole RGB space, we repeated the staircase procedure for three different pairs of modulating phosphors (R-G, B-G, and B-R) and estimated the three corresponding ratios of phosphor intensities (R/G, B/G, B/R).

To account for changes in isoluminance as a function of eccentricity, we measured RGB scale factors at three concentric nonoverlapping annuli (2.167° width), spanning the spatial extent of our stimulus (1.5–8° in radius). For each subject, we made three independent measures of the three ratios (R/G, B/G, B/R) at each of the three eccentricities (for a total of 27 measurements) and used the corresponding averages as our estimate of the isoluminance ratios. The RGB scale factors did not change appreciably with eccentricity; nonetheless, we used the ratios measured for the three eccentricities to derive a first-order (linear) estimate of the RGB scale factor as a function of eccentricity, which we used to adjust the color of dots at different eccentricities. Each dot's color was further jittered by a small amount of luminance noise to counteract any small hue-dependent luminance bias that our isoluminance measurements failed to account for. The luminance noise consisted of white noise (SD of 5.5 cd/m^2, low-passed filtered by a Gaussian kernel in the frequency domain with SD of 0.5 cycles/s). Finally, we used the gamma curves for each phosphor to ensure the color outputs had the desired intensity.

Motion specifications. For each surface, the direction of motion was specified by the coherent translation of individual dots in a specific direction between successive monitor frames. The direction of motion of the two surfaces changed smoothly around the clock with temporal periods of 25.25 s (T_{m1}) and 21.82 s (T_{m2}), corresponding to 22 and 25 full rotations over the 8-min scan. The dynamics of the direction of motion in the two surfaces can be mathematically formulated as follows:

$$\theta_i(t) = 2\pi \times \frac{t}{T_{mi}} + \phi_i + n \qquad (2)$$

where θ corresponds to the direction of motion in radians, t corresponds to the elapsed time in seconds, i indexes each surface, ϕ is a phase parameter that determines the surface's initial motion direction, and n corresponds to the dispersion added during a motion event (see *Motion events*).

The color and direction of motion of the dots were subject to a number of constraints. We use the terms "birth," "death," "age," and "lifetime" to refer to the beginning, the end, the number of elapsed frames since birth, and the duration (in frames) of coherent translation of a dot, respectively. At stimulus onset ($t = 0$) each dot was assigned a random age from 0 to 11. The initial color and motion of each dot

was specified by setting $t = 0$ (*Eqs. 1* and *2*). After each monitor refresh, the age of every dot was incremented by 1. Each dot maintained its color, its luminance, and its direction of motion for the duration of its lifetime, which was fixed to 12 frames (200 ms). If a dot moved outside the annulus during its lifetime, then its position was wrapped around to the other side of the annulus. The death of each dot (at an age of 12 frames) led to the birth of a new dot (age = 0) at a random position within the annulus. The color and direction of motion of each newly born dot were adjusted based on *Eqs. 1* and *2* at the new time t (12 frames past the previous birth). Because the dots' ages at stimulus onset were assigned randomly, at any given moment in time each surface maintained a small dispersion around its average color and direction of motion.

Stimulus events. To provide a task for the subjects, each surface was subject to brief perturbations in color and motion direction, which we refer to as "color events" and "motion events," respectively. Each event type (color and motion) in each of the two surfaces occurred independently with an average frequency of once per 7 s. Each event lasted 1 s and was followed by an absolute refractory period of 1 s, during which events never occurred. After the refractory period, the probability of a subsequent event was constant (e.g., a flat hazard rate). The stimulus events, which are described in more detail below, can be seen in the Supplementary Movie 1 by tracking one surface as it evolves over the course of the movie.

Color events. A color event was characterized by a transient increase in the variance of the color in a surface, which lasted a total of 1 s. To modulate the variance, we added noise to the color of individual dots that were born during the event. The additional noise was controlled by a random variable, n, with uniform distribution, that was added to the phase of a* and b* (*Eq. 1*). In the absence of a color event, the value of n was set to 0 (i.e., no additional variance). During a color event, the upper limit in the range of n increased linearly from 0 to n_{max} for the first 0.5 s of the event and then decreased linearly back to 0 for the second half. For each dot, n_{max} was specified by a random draw between $-2\pi/3$ and $2\pi/3$ radians.

Motion events. Similar to the color event, when assigning a direction of motion to a dot born during a motion event, a random variable, n (*Eq. 2*), was added to the phase of the direction of motion. The value of n was set to 0 when no motion event was present. During a motion event, the upper limit in the range of n increased linearly from 0 to n_{max} for the first 0.5 s of the event and then decreased linearly back to 0 for the second half. For each dot, n_{max} was specified by a random draw between $-2\pi/3$ and $2\pi/3$ radians.

fMRI scanning sequence. The stimulus was identical across scans. Before a scan, subjects were cued to track one of the four surface features or to perform a demanding task at fixation. Subjects where scanned over two separate scanning sessions, held on separate days. The sequence of tasks (i.e., conditions) in a typical scanning day was as follows: *1*) perform the fixation task, *2*) track the motion of *surface 1*, *3*) track the motion of *surface 2*, *4*) repeat the fixation task, *5*) track the motion of *surface 2*, and *6*) track the motion of *surface 1*. For half of the subjects, the motion task was performed on *day 1* and the color task on *day 2*; for the other half, this sequence was reversed.

Behavioral task: fMRI. For each scan, observers were cued to track a single surface and to detect events within a single feature, either the motion or color of that surface. We refer to this event as the target event and to the other three event types as distractor events. For example, when subjects were instructed to detect events in the color of *surface 1*, the color events in *surface 1* were the target events and the motion events in *surface 1* as well as both color and motion events in *surface 2* were distractor events. Observers were instructed to press a response button immediately after detecting a target event, while ignoring all distractor events.

Behavioral analysis: fMRI. Responses were divided into *1*) hits, *2*) false alarms, and *3*) selection errors (misses were separately tallied). A response was classified as a hit if the subject pressed the response button within a 2-s window after the onset of the target

event. For each scan, the hit rate was computed by dividing the number of hits by the total number of target events. Target events that were not followed by a button press were referred to as misses. A response was classified as a false alarm if no event (target or distractor) preceded the response within a 2-s window. Finally, a response was classified as a distractor response if the 2 s preceding the response contained any distractor event but no target event.

Psychophysics experiment. On each trial subjects were presented with a 5-s stimulus, similar to the one we used in the fMRI experiment. A random initial color (ϕ_1) and motion direction (θ_1) was selected for the first surface on every trial. The initial color and motion direction of the second surface was shifted by 90° in color and motion space from the first surface; i.e., $\phi_2 = \phi_1 + \pi/2$ (*Eq. 1*), and $\theta_2 = \theta_1 + \pi/2$ (*Eq. 2*).

There were two single-task conditions and one dual-task condition. Subjects were cued to track the color, the motion, or both color and motion within one of the surfaces. At 300 ms poststimulus onset, the color and/or the motion of one surface was cued. To cue motion, the speed of the cued surface increased; to cue color, the luminance of the cued surface increased. In both cases the intensity change was stepwise over the duration of the cue (300 ms).

Color and motion events occurred with 50% probability within each feature. The onset of an event occurred randomly, and with equal probability, between 1 and 4 s from stimulus onset. To prevent subjects from using a switching strategy in the dual-task condition, on trials when both a color and motion event occurred within the same surface, they were constrained to occur simultaneously. Thus, on 25% of the trials, the motion and color events co-occurred. After stimulus offset, subjects reported (via the key press) whether or not there was any target event in the stimulus. The yes/no responses for each task were mapped to separate keys, one set for each hand. The response order was counterbalanced across subjects in the dual-task condition.

Subjects performed each condition in blocks of 30 trials. The order of the blocks was counterbalanced between subjects. Subjects practiced the task over one or two 1-h sessions. After practice, each subject ran four blocks of 30 trials for each condition, for a total of 120 trials per condition.

Retinotopic mapping procedure. Retinotopic mapping was obtained in a single 1-h session by using a flickering checkerboard restricted to a rotating wedge, an expanding annulus, and an alternating pair of wedges covering the vertical and horizontal meridian [stimulus flicker 6 Hz, and wedges subtended 40° of polar angle (Engel et al. 1994; Sereno et al. 1995)]. With this procedure, visual areas V1, V2v, V2d, V3v, V3d, and hV4 were drawn by hand on the inflated representation of the cortical surface using BrainVoyager QX (version 1.9.10; Brain Innovation, Maastricht, The Netherlands). Ventral and dorsal areas were collapsed together for the analysis. Ventral area hV4 was defined to include an entire hemifield representation (Wandell et al. 2005). A functional localizer was used during each experimental session to define motion-selective area MT+.

fMRI data acquisition and analysis. MRI scanning was performed on a Phillips Achieva 3-Tesla scanner, located at the University of Washington Magnetic Resonance Research Laboratory, equipped with an 8-channel head coil. Anatomic T1-weighted images were acquired at $1 \times 1 \times 1$-mm resolution. Whole brain, 32-transverse slice functional images were acquired at $3.438 \times 3.438 \times 3.5$-mm resolution (repetition time, 2,000 ms; echo time, 30 ms; flip angle, 76°; scan resolution, 64×64; field of view, 220 mm; slice thickness, 3.5 mm; no gap). No smoothing was applied during preprocessing.

Each scan was motion corrected using BrainVoyager QX. Experimental scans were coregistered to the anatomic retinotopy scans. When a functional data set was coregistered to a higher-resolution anatomic data set, the time course of each functional voxel was assigned to all of the anatomic voxels that fell within the boundaries of the functional voxel. All redundant time courses that were created by upsampling were discarded.

Region of interest selection. Localizer scans were run at the beginning and end of each experimental session. A general linear model (GLM) was used to find voxels that responded strongly to the region of visual space corresponding to the extent of the stimulus. Regressors were created in the GLM by convolving a gamma function with the boxcar stimulus protocol. The functional localizer consisted of 20-s blocks of fully coherent moving achromatic dots (randomly reassigned 1 of 8 possible directions of motion every second, 200-ms limited lifetime), static dots (redrawn in a random configuration every second), and a blank screen. All other properties of the localizer were set to match the experimental configuration, including the dimensions of the stimulus aperture, dot size, speed, density, etc. We selected voxels in V1, V2, V3, and hV4 that responded more strongly to the motion than to the blank condition. We defined MT+ as a contiguous patch of medial temporal cortex that responded more to the motion condition than the static condition ($P < 0.05$, Bonferroni corrected for multiple comparison).

Frequency analysis. Over the course of an 8-min scan, the color of *surfaces 1* and *2* completed 19 and 28 full cycles through color space, while the motion direction of *surfaces 1* and *2* completed 25 and 22 full cycles. We quantified the periodicity of the hemodynamic response at the frequencies associated with the color and direction of motion of the two surfaces from the amplitude of the corresponding fundamental harmonics in the Fourier spectrum. Hereafter, we will use the term "harmonic response" to refer to the fundamental (first) harmonic response. We measured the Fourier spectrum of responses in different visual areas by applying MATLAB's fast-Fourier transform to the average time course of the BOLD signal in those areas. We analyzed the frequency spectrum in different visual areas and compared the amplitude of the harmonic responses associated with color and direction of motion of the two surfaces across attention conditions.

Simulations. To demonstrate our frequency-tagging approach, we performed a simulation of 200 feature-selective voxels. The purpose of this simulation was to demonstrate how a stimulus response can be extracted from the frequency spectrum averaged over a population of voxels with random feature-tuning profiles. Each voxel consisted of 16 feature-selective channels : 8 direction-selective channels and 8 color-selective channels evenly spaced between 0° and 360°. We modeled the response of each channel to each cyclically modulated stimulus feature by a sine wave at the corresponding frequency whose phase was determined by the channel's feature preference. Each channel's response to the two superimposed surfaces was then measured as the linear sum of the four sine waves associated with the four stimulus features (2 directions of motion and 2 colors).

To simulate the random biases in each voxel's direction and color tuning (Fig. 2, *A* and *B*), each of the underlying 16 channels was assigned a random weight between 0 and 1, and the voxel's population response was computed as the sum of the 16 sine waves, each scaled by its corresponding weight. To simulate fMRI signal, each voxel's population response was convolved with a standard hemodynamic impulse response function (gamma function: $n = 3$, $\tau = 1.5$, delay = 2 s). Finally, we normalized the time course by subtracting and then dividing the time course by its mean, and added white noise (0 mean, 0.1 SD) to simulate the BOLD signal (see Fig. 2*C*). To quantify the harmonic response across the population of 200 simulated voxels, we performed a fast-Fourier transformation of each voxel's time course and then averaged across the resulting amplitude spectrum (see Fig. 2*D*).

The effect of feature-based attention was modeled by amplifying the sinusoid corresponding to the attended feature by a factor of 1.2 while attenuating the sinusoids corresponding to other features by a factor of 0.83 (see Fig. 2, *E* and *F*). The effect of object-based attention was modeled by amplifying both sinusoids that correspond to the attended surface by a factor of 1.2 while attenuating the sinusoids corresponding to the other surface by a factor of 0.83 (see Fig. 2, *G* and *H*). The gain factors were chosen for demonstration purposes and were not constrained by the data.
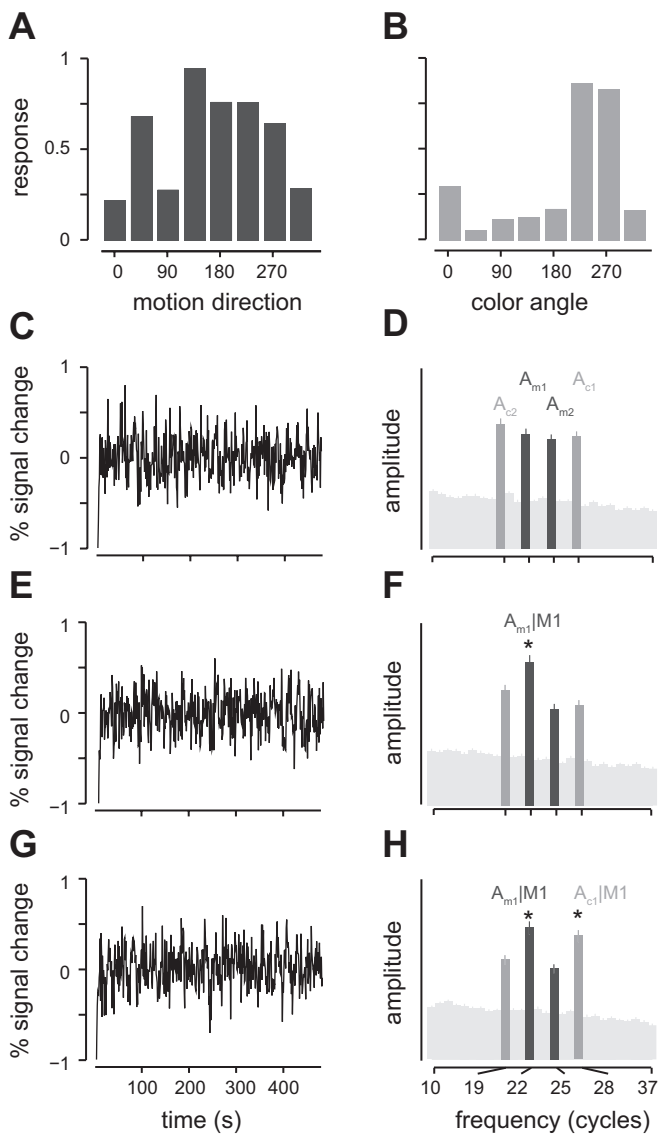
Fig. 2. Simulated harmonic response of 200 hypothetical voxels. *A* and *B*: tuning for the direction of motion and color, respectively, for an example voxel. *C*: time course of the functional magnetic resonance imaging (fMRI) signal for an example voxel in response to the stimulus (Fig. 1). *D*: average Fourier amplitude spectrum across the 200 simulated voxels. Error bars correspond to 1 SE. The spectrum is plotted for a subset of the harmonics between 10 and 37 cycles/scan that span the 4 stimulus components. Black bars at 22 and 25 cycles/scan correspond to the amplitude of the motion harmonic for *surface 1* (A_{m1}) and *surface 2* (A_{m2}). Gray bars at 19 and 28 cycles/scan correspond to the amplitude of the color harmonic for *surface 1* (A_{c1}) and *surface 2* (A_{c2}). *E*: time course of response for an example voxel subject to the effect of feature-based attention toward the direction of motion in *surface 1* (M1). *F*: average amplitude spectrum across the population of voxels demonstrating an enhanced response to M1 (A_{m1}|M1, indicated by asterisk) relative to the other stimulus frequencies. *G*: time course of response for an example voxel subject to the effect of object-based attention when subject attended M1. *H*: average spectrum across the voxels showing the effect of object-based attention enhancing the harmonic responses to both direction of motion and color of *surface 1* (A_{m1}|M1 and A_{c1}|M1, indicated by asterisks) relative to the response to the unattended surface.

RESULTS

Subjects were scanned while viewing a stimulus consisting of two superimposed surfaces composed of dot fields with unique color and motion conjunction (Fig. 1*A*). The direction of motion and the color of each surface slowly changed in a

periodic fashion with a unique temporal period for each feature (Fig. 1, *B* and *C*, and Supplementary Movie 1). Before a scan, subjects were cued to track the motion or color of one of the two surfaces and performed an ongoing task in which they were instructed to respond with a button press every time they detected a target event in the cued surface feature (while ignoring distractor events). Events were defined as brief dispersions in the motion or color coherence (see METHODS, *Stimulus events*).

Behavioral results during fMRI data acquisition. All six subjects were able to track the cued surface feature to respond to target events and ignore distractor events. Table 1 shows the proportion of button presses following target events (target response), distractor events (distractor response), and no events (false alarms) for each subject. The majority of button presses were associated with a target event (0.91 ± 0.04, mean $\pm$ SD). There were relatively fewer distractor responses (0.07 ± 0.04) and very few false alarms (0.02 ± 0.02). The relative high rate of target responses suggests that subjects were able to track the cued surface feature as it progressed through feature space. We intentionally made the magnitude of the event transients small to reduce distractor interference. This resulted in a difficult detection task. Thus, even though most responses were to targets, subjects missed $44 \pm 6\%$ of target on average.

The harmonic hemodynamic response. Previous work showed that individual voxels in different visual areas could exhibit weak but reliable selectivity for stimulus features, including color and direction of motion (Brouwer and Heeger 2009; Kamitani and Tong 2006). Consequently, by changing the color and direction of motion of the stimulus in a circular fashion, we should be able to modulate the response of color- and direction-selective voxels in a periodic fashion. Because the BOLD signal is sluggish, we used relatively low frequencies (long periods) over which to modulate the stimulus features: 25 and 22 cycles/scan (8 min/scan) for the direction of motion and 19 and 28 cycles/scan for the color of *surfaces 1* and *2*, respectively. We hypothesized that attention modulates the gain of feature-selective neurons (Martinez-Trujillo and Treue 2004), leading to a change in the amplitude of the corresponding harmonics in the hemodynamic response (Boynton 2005b). We therefore analyzed the frequency spectrum from the BOLD time course in different visual areas and compared the amplitude of the harmonic responses associated with color and direction of motion of the two surfaces across attention conditions.

To facilitate the presentation of results, we developed nomenclature to refer to the different attention conditions and

Table 1. *Summary of behavioral responses for each subject in the detection task in the scanner*

	Subject					
	1	*2*	*3*	*4*	*5*	*6*
Target response	0.86	0.96	0.91	0.90	0.89	0.96
Distractor response	0.13	0.04	0.07	0.09	0.05	0.03
False alarm	0.01	0.00	0.02	0.01	0.06	0.01

The 3 response rates sum to 1.0 and are defined as follows: target response, proportion of responses following a target event; distractor response, proportion of responses following a distractor event; false alarm, proportion of responses following no target or distractor event.

different response harmonics: A_{m1} and A_{m2} are the amplitudes of the harmonic responses associated with the motion feature in *surfaces 1* and *2*, respectively; A_{c1} and A_{c2} are the amplitudes of the harmonic responses associated with the color feature in *surfaces 1* and *2*, respectively; M1 and M2 refer to conditions in which subjects tracked the motion of *surfaces 1* and *2*, respectively; and C1 and C2 refer to conditions in which subjects tracked the color of *surfaces 1* and *2*, respectively.

To demonstrate the analysis and the competing feature- vs. object-based predictions, we simulated the response of 200 voxels with random motion and color biases (Fig. 2). To do so, we modeled each voxel as comprising 16 feature-selective channels: 8 for motion and 8 for color. Each channel was assigned a random weight so that the population response exhibited randomness for both direction of motion (Fig. 2*A*) and color (Fig. 2*B*). We simulated each voxel's response by the sum of the responses of the underlying channels plus noise (Fig. 2*C*). Averaging the amplitude spectrum across all 200 simulated voxels reveals the harmonic response to each stimulus component (Fig. 2*D*). The amplitudes labeled A_{m1} and A_{m2}, at 22 and 25 cycles/scan, refer to the motion harmonics to *surfaces 1* and *2*, respectively, and A_{c1} and A_{c2}, at 28 and 19 cycles/scan, refer to the corresponding color harmonics. The feature-based hypothesis predicts that attention will modulate the response to the cued feature. The object-based hypothesis predicts that attention will modulate the response to the cued feature and to the other feature of the same surface. We modeled these two alternatives by appropriate gain changes in the response of the underlying channels (see METHODS, *Simulation*). Figure 2, *E* and *F*, shows the time course of the response and the corresponding amplitude spectrum when feature-based attention was directed to the motion of surface 1 (M1). As expected, feature-based attention enhances the amplitude of the motion harmonic in *surface 1* (A_{m1}|M1) relative to the other stimulus components. Figure 2*G* shows the time course associated with the effect of object-based attention. In

this case, amplitudes associated with both the motion and color of *surface 1* (A_{m1}|M1 and A_{c1}|M1) are enhanced.

We used a similar methodology to analyze the data collected from the scanning sessions. We collapsed the responses across all voxels within each visual area and then averaged across subjects to make an overall qualitative assessment of the effects of attention on harmonic responses (Figs. 3–7). The between-subjects averaged amplitude spectrum for V1 is shown in Fig. 3 for each attention condition.

Figure 3*A* shows the amplitude spectrum when the motion of *surface 1* was tracked (M1). In this condition, A_{m1}|M1 was weakly enhanced relative to the neighboring noise response, suggesting that attention to M1 could enhance the response to the corresponding feature. This qualitative enhancement was not present for the motion in *surface 2* (A_{m2}|M1), suggesting that our observation is not due to spatial attention or nonspecific enhancement of direction-selective mechanisms in V1. When subjects tracked the motion of *surface 2* (M2), A_{m2}|M2 appears to have been enhanced, rising above the surrounding noise relative to A_{m2}|M1, and A_{m1}|M2 appears to have been reduced in amplitude compared with A_{m1}|M1 (Fig. 3*B*). These results indicate that, in V1, the effect of feature-based attention to the direction of motion is surface specific.

When subjects attended the color of either surface, the color harmonics in V1 (A_{c1}|C1 and A_{c2}|C2) do not appear to have risen above the surrounding noise (Fig. 3, *C* and *D*). Surprisingly, however, attention to color of a given surface seems to have enhanced the responses associated with the motion of that surface: in the C1 and C2 conditions, respectively, A_{m1}|C1 (Fig. 3*C*) and A_{m2}|C2 (Fig. 3*D*) appears to be greater than the noise harmonics. In fact, the overall pattern of amplitude responses across the four stimulus frequencies looks similar regardless of whether the subjects track the motion or color of a surface (Fig. 3, *A* vs. *C*, and *B* vs. *D*). These results suggest that the observed modulations in V1 BOLD responses are associated with surface-specific attentional selection for the direction of motion.
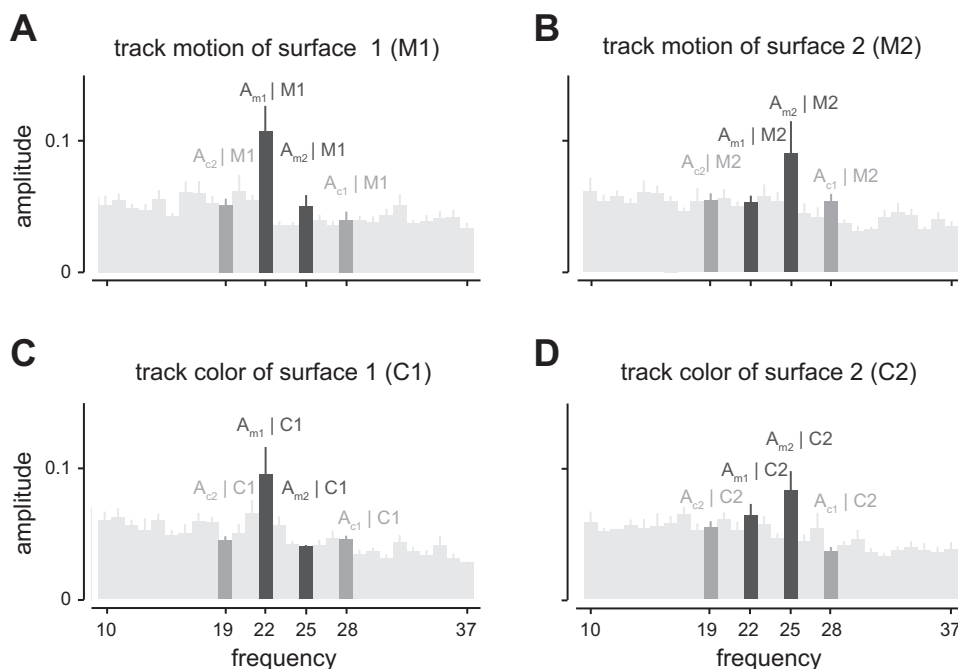

Fig. 3. Average amplitude spectrum of V1 hemodynamic responses under each attention condition. Each panel shows the amplitude spectrum of the hemodynamic responses averaged across the 6 subjects. Error bars show SE of the mean across the 6 subjects. *A*: amplitude spectrum when subjects were cued to track the motion of *surface 1* (M1). The amplitude of the harmonic response associated with the motion of *surface 1* shown in black (A_{m1}|M1 at 22 cycles/scan) was higher than the harmonic response associated with the motion of *surface 2* (A_{m2}|M1 at 25 cycles/scan). In this condition, the harmonic responses to the color, shown in dark gray (A_{c1}|M1 at 28 cycles/scan for *surface 1* and A_{c2}|M1 at 19 cycles/scan for *surface 2*), are comparable to the response amplitude at nonstimulus harmonics (gray). *B*: same as *A* for the condition in which subjects were cued to track the motion of *surface 2* (M2). *C* and *D* show similar measurements for when subjects tracked the color of *surfaces 1* (*C1*) and *2* (*C2*), respectively.
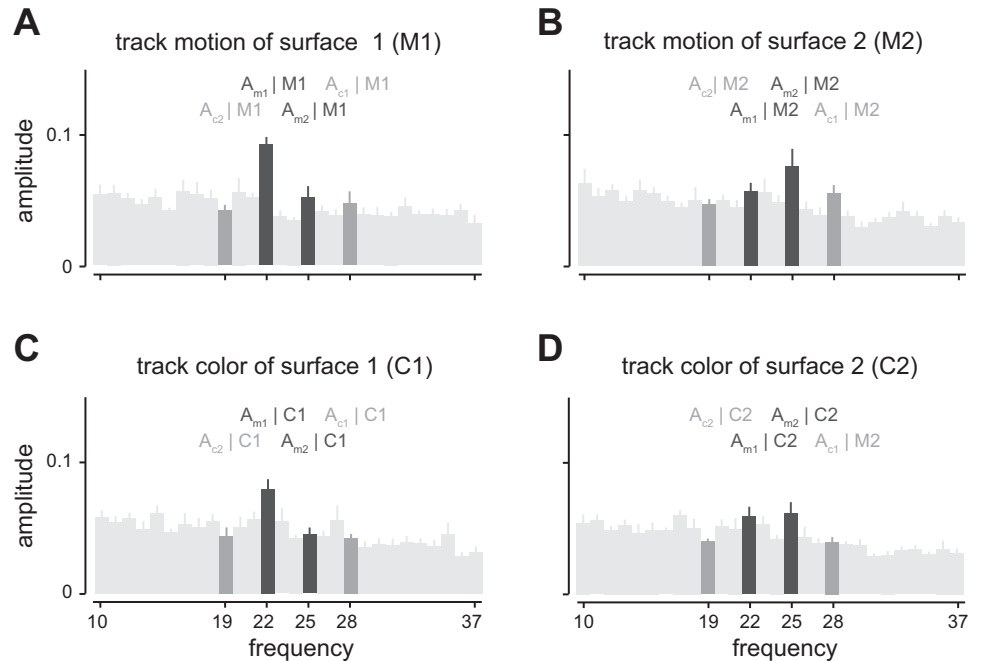
Fig. 4. Average amplitude spectrum of V2 hemodynamic responses under each attention condition. All the details are the same as in Fig. 3.

A qualitatively similar pattern of responses was observed in visual areas V2 and V3 (Figs. 4 and 5). Area V4 followed the same trend but was less reliable than the earlier visual areas (Fig. 6). Area MT+ did not produce a reliable harmonic response to our stimulus (Fig. 7).

In a separate experiment, we attempted to measure a baseline response to the stimulus while the subject's attention was diverted by a fixation task. Under this attentional condition the stimulus failed to drive voxels at the stimulus harmonics across all visual area (data not shown). Attention to the stimulus was therefore required to drive the harmonic response above the noise.

Feature-based attention. The qualitative pattern of selective enhancement of A_{m1} and A_{m2} in M1 and M2 conditions,

respectively, corresponds to the effects of feature-based attention on the motion response (Fig. 3, *A* and *B*). We developed a feature-based attention index (FI) to quantify the magnitude of the feature-based effect on the motion and color response. The index measures the relative change in the amplitude of the harmonic response to a given feature when it is attended vs. when it is unattended. For example, FI_{m1} is the normalized difference between the amplitude of the harmonic response to the motion of *surface 1* (A_{m1}) under two different attention conditions, M1 and M2.

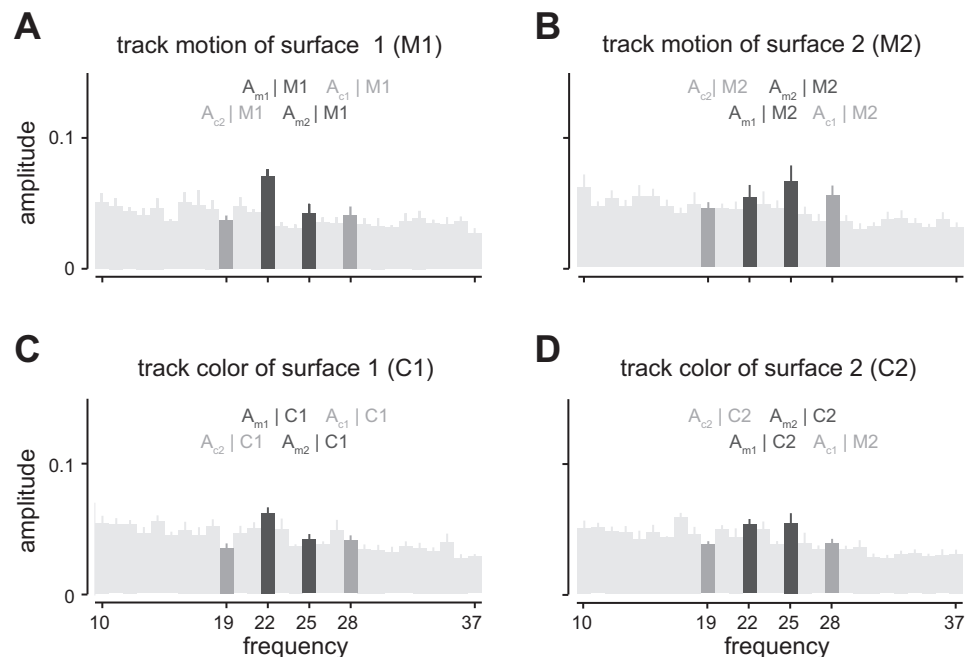$$FI_{m1} = \frac{A_{m1}|M1 - A_{m1}|M2}{A_{m1}|M1 + A_{m1}|M2} \qquad (3)$$

Fig. 5. Average amplitude spectrum of V3 hemodynamic responses under each attention condition. All the details are the same as in Fig. 3.

A track motion of surface 1 (M1)

B track motion of surface 2 (M2)

C track color of surface 1 (C1)
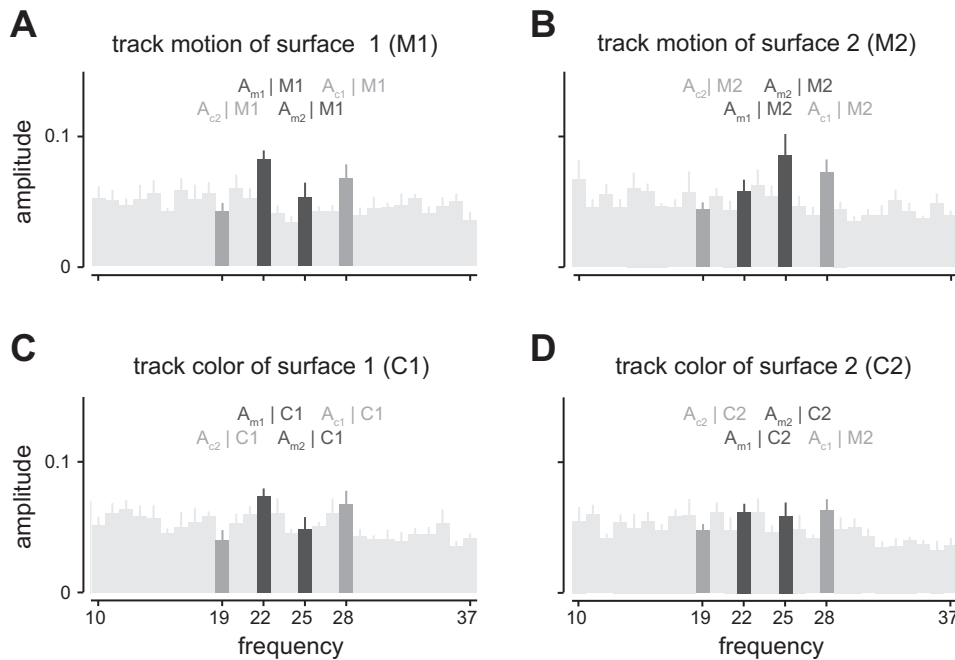
D track color of surface 2 (C2)

Fig. 6. Average amplitude spectrum of V4 hemodynamic responses under each attention condition. All the details are the same as in Fig. 3.

This formulation of the FI provides a simple metric for feature attention that ranges between -1 and 1, with 0 corresponding to no attentional effect and 1 and -1 corresponding to strong enhancement and suppression of responses. Figure 8, *A* and *B*, shows the FI for the motion response (FI_{m1} and FI_{m2} for *surfaces 1* and *2*, respectively) for each visual area. For both surfaces, the FI was positive in V1, V2, V3, and V4, but the effect was weak and only significant in areas V1 and V2 for one of the surfaces [$t(6) = 3.25$ for V1 and 2.80 for V2, $P < 0.05$]. The effect was weakest in area MT+ where direction selectivity is strong (Tootell et al. 1995; Zeki et al. 1991).

In many cases, FI was not significant but showed a positive trend. To determine whether the weak positive trend in FI was reliable, we asked whether a similar trend was evident in both hemispheres. We first averaged FI_{m1} and FI_{m2} to compute a single index (FI_{mot}) for each visual area (Fig. 8*C*) and then compared this value between the two hemispheres for the five visual areas (V1, V2, V3, V4, and MT+) in the six subjects. This analysis (Fig. 8*C*, *inset*) showed that, across areas and subjects, FI_{mot} was significantly correlated between the two hemispheres [$r(30) = 0.85$, $P < 0.001$], suggesting that the feature-based attention effect for motion was reliable.

We defined a similar feature index to quantify the effect of feature-based attention on the color response in each visual area. Figure 8, *D* and *E*, shows the FI for the color of the two surfaces (FI_{c1} for *surface 1* and FI_{c2} for *surface 2*) for each

A track motion of surface 1 (M1)

B track motion of surface 2 (M2)

C track color of surface 1 (C1)

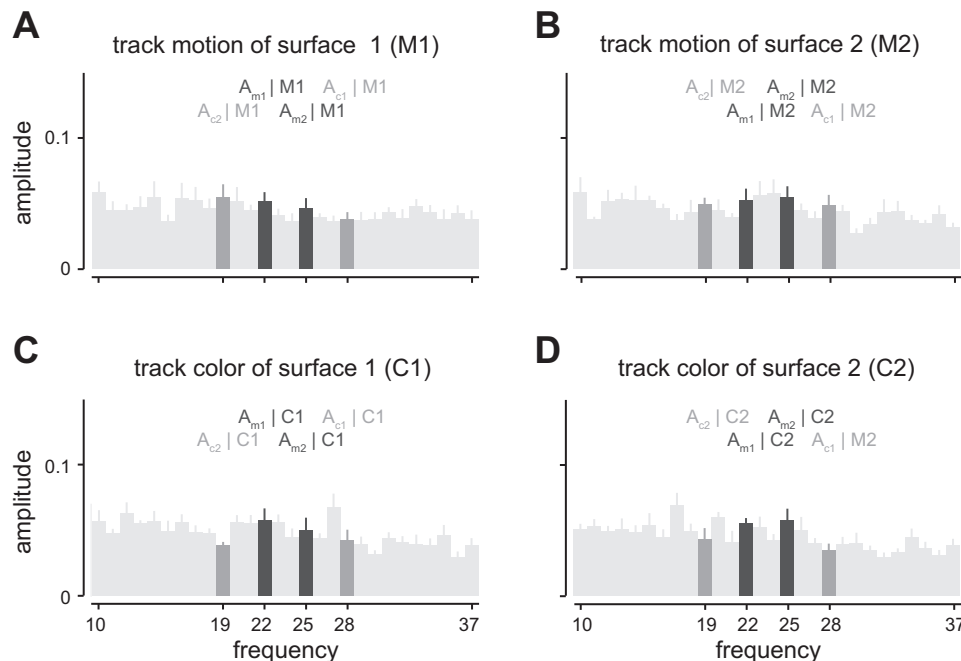**D** track color of surface 2 (C2)

Fig. 7. Average amplitude spectrum of MT hemodynamic responses under each attention condition. All the details are the same as in Fig. 3.

Fig. 8. Feature-based and object-based attention indexes across conditions and visual areas. Each panel shows the value of attention index averages across the 6 subjects for areas V1, V2, V3, V4, and MT+. *A–F*: feature-based attention indexes quantify the effect of attention on the amplitude of the motion harmonics (*A–C*) and the effect of attention on the color harmonics (*D–F*). *A–C*: feature-based attention index for the direction of motion of *surface 1* (FI_{m1}), the direction of motion of *surface 2* (FI_{m1}), and the overall average of the 2 surfaces (FI_{mot}). *D–F*: feature-based attention index for the color of *surface 1* (FI_{m1}), the color of *surface 2* (FI_{m1}), and the overall average of the 2 surfaces (FI_{mot}). *G–L*: object-based attention indexes quantify the effect of attention on the amplitude of the color harmonics (*G–I*) and the effect of attention on the motion harmonics (*J–L*). *G–I*: object-based attention index for the direction of motion of *surface 1* (OI_{m1}), the direction of motion of *surface 2* (OI_{m1}), and the overall average of the 2 surfaces (OI_{mot}). *J–L*: object-based attention index for the color of *surface 1* (OI_{m1}), the color of *surface 2* (OI_{m1}), and the overall average of the 2 surfaces (OI_{mot}). In each row, the *inset* shows the overall average attention index computed from hemodynamic responses in the right hemisphere (ordinate) and left hemisphere (abscissa) for each subject and each visual area. Error bars show SE of the mean across the 6 subjects. Asterisks denote the attention index values that are significantly different from zero ($P < 0.05$, 1-sample *t*-test).
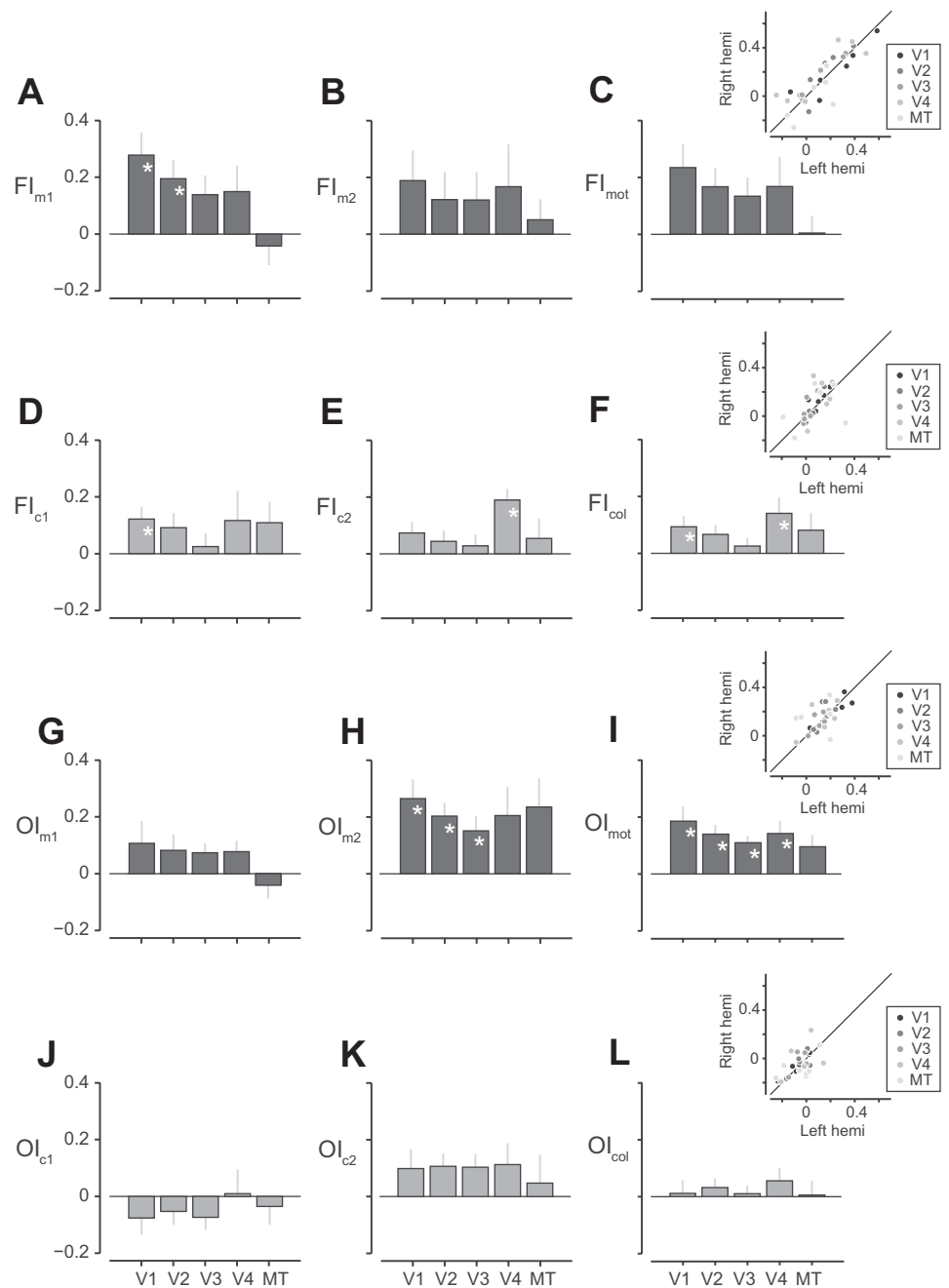
visual area. The FI was positive across all visual areas for both surfaces, but the effect was only significant in V1 for *surface 1* and in V4 for *surface 2* [$t(6) = 2.80$ for V1 and 4.88 for V4, $P < 0.05$].

We examined whether the effect of feature-based attention to color was reliable by comparing the FI values between hemispheres. To do so, we averaged FI_{c1} and FI_{c2} to compute a single index (FI_{col}) that quantified the effect of feature-based attention for color (Fig. 8*F*). Overall, there was a positive effect of feature-based attention on the color response across visual areas, which reached significance in areas V1 and V4 [$t(6) = 2.36$ for V1 and 1.34 for V4, $P < 0.05$]. We then compared FI_{col} in the left and right hemispheres across five visual areas (V1, V2, V3, V4, and MT+) in the six subjects (Fig. 8*F*, *inset*) and found that FI_{col} was

significantly correlated between the two hemispheres [$r(30) = 0.53$, $P < 0.01$], suggesting that the index is reliable.

Object-based spread of attention. The selective enhancement of the motion-driven components A_{m1} and A_{m2} in conditions where color was attended (C1 and C2) suggests that attention to the color of each surface enhanced the motion-related signals of the attended surface (Fig. 3, *C* and *D*). To quantify this spread of attention to the task-irrelevant feature of a surface, we developed an object-based attention index (OI) to quantify the magnitude of the object-based effect on both the motion and color response. The index measures the relative change in the amplitude of the harmonic response to a given feature when the other feature of that same surface is attended versus when the other surface is attended. For example, OI_{m1}

is the normalized difference between A_{m1} under two different attention conditions, C1 and C2.

$$OI_{m1} = \frac{A_{m1}\big|C1 - A_{m1}\big|C2}{A_{m1}\big|C1 + A_{m1}\big|C2} \qquad (4)$$

Like the FI, the OI ranges between -1 and 1. Figure 8, G and H, shows the OI for the motion response of each surface (OI_{m1} and OI_{m2} for *surfaces 1* and *2*, respectively) for each visual area. OI_{m2} was positive in V1, V2, V3, and V4 for both surfaces and was significantly greater than zero in V1, V2, and V3 for *surface 2* [$t(6) = 3.87$, 4.39, and 2.83, respectively, $P < 0.05$].

To summarize the effect of object-based attention on the motion response, we averaged OI_{m1} and OI_{m2} together to form OI_{mot} (Fig. 8I). OI_{mot} was positive in all five visual areas and was significant in V1, V2, V3, and V4 [$t(6) = 3.43$, 4.34, 4.66, and 2.82, respectively, $P < 0.05$]. To test the reliability of this statistic, we compared OI_{mot} between hemispheres (Fig. 8I, *inset*) across all five visual areas. OI_{mot} was significantly positively correlated between hemispheres [$r(30) = 0.54$, $P < 0.01$], suggesting that it is a reliable index.

Finally, we quantified the effect of object-based attention to the color harmonics under conditions in which the motion of each surface was cued. For these conditions, we found no consistent object-based effect; OI_{c1} and OI_{c2} were not significantly different from zero for any of the visual areas we tested. OI_{c1} was, on average, negative across visual areas (except V4) and OI_{c2} was positive (Fig. 8, J and K). The average effect (OI_{col}) across the two surfaces was positive but not significant (Fig. 8L), although the OI_{col} values were significantly correlated between the two hemispheres [Fig. 8L, *inset*; $r(30) = 0.58$, $P < 0.001$].

No dual-task cost when tracking two features within a surface. Our fMRI results suggest that attention spreads to the uncued surface feature. If attention is deployed to both surface features regardless of which feature is cued, then the observer may have equal perceptual access to both surface features. This hypothesis can be tested by comparing behavioral performance in a single-task condition, identical to the task in our fMRI experiment, with behavioral performance a dual-task condition in which the observer is instructed to monitor both surface features simultaneously.

We conducted a separate psychophysical experiment outside the scanner to assess the behavioral consequences of our observed fMRI results. Subjects performed a yes-no detection task while viewing short 5-s segments of the same stimulus used in the fMRI experiment. In the single-task condition, subjects were either cued to track the motion or the color of one of the two surfaces. In the dual-task condition, subjects were cued to track both features within one of the two surfaces (see METHODS, *Psychophysics experiment*).

We compared performance between the single- and dual-task conditions to see whether dividing attention across features within a surface would result in a cost in behavioral performance. As evidenced by the scatter plot in Fig. 9, performance of individual subjects on the single- and dual-task conditions was comparable, with no significant difference across subjects [single-task minus dual-task: $t(6) = 0.62$, $P = 0.56$ for motion, and $t(6) = 0.20$, $P = 0.85$ for color].
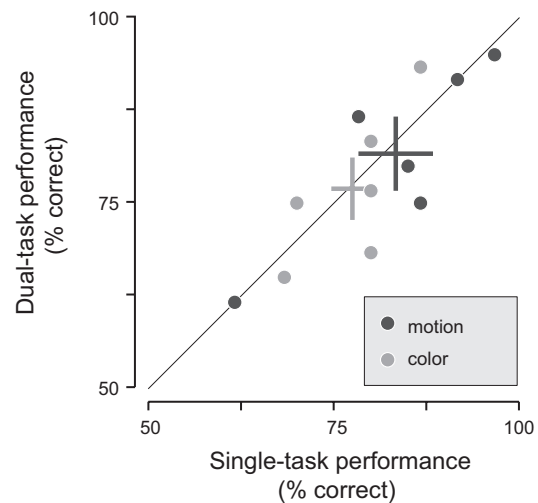

Fig. 9. Behavioral effect of object-based attention. Each subject's single-task performance (abscissa) is plotted against their dual-task performance (ordinate) for the motion task (dark gray points) and color task (light gray points). The identity line indicates equal performance on both single- and dual-task conditions. The black and gray crosses correspond to the average performance across the 6 subjects for the motion and color task, respectively. The horizontal and vertical extent of each cross corresponds to the SE of the mean for the single- and dual-task performance levels.

DISCUSSION

Both behavioral and physiological evidence has provided support for object-based attention. Theories of object-based attention posit that all features of a behaviorally relevant object are selected. Behavioral studies support this claim, showing that attention can be divided across multiple features within an object without a reduction in performance, whereas selecting features from different objects is much more difficult (Blaser et al. 2000; Bonnel and Prinzmetal 1998; Duncan 1984; Ernst et al. 2012).

In an early neuroimaging study of object-based attention, O'Craven et al. (1999) presented subjects with either a moving face superimposed on a stationary house or a moving house superimposed on a stationary face. They found that when subjects attended to one attribute of the moving face (either the identity of the face or the direction of motion), the amplitude of the BOLD response in both the motion-sensitive area MT+ (MT/MST) and the fusiform face area was enhanced. Similarly, attending attributes of the moving house enhanced responses in MT/MST as well as the parahippocampal place area. This finding provides evidence that object-based attention enhances the response of brain areas representing the various attributes of a relevant object.

However, region-of-interest-based analyses that rely on an averaged BOLD signal, like those used by O'Craven et al. (1999), cannot specify whether object-based attention operates at the level of functional areas (e.g., area MT/MST) or at the finer level of feature-selective mechanisms within an area (e.g., direction-selective neurons within MT). To overcome this limitation, several studies have exploited the inhomogeneities in sensory representations to decode feature-based attentional modulations from the pattern of hemodynamic responses (Brouwer and Heeger 2009; Haynes and Rees 2005; Kamitani and Tong 2005, 2006; Serences and Boynton 2007). For example, pattern classification has been used to predict which of two superimposed surfaces a subject attended (Kamitani and

Tong 2005). Pattern classification has also been used to classify object identity in higher visual areas (Grill-Spector and Sayres 2008; Haxby et al. 2001; O'Toole et al. 2005) and to decode the effects of selective attention on the representation of object-based information in extrastriate cortex (Chen et al. 2012). Successful decoding provides evidence that selection of a feature could bias the pattern of responses across voxels toward the pattern produced by the attended feature in isolation (Boynton 2005a). Here, we developed a frequency-tagging scheme that exploits these biases to assess the correlate of object-based attention at the level of feature-selective mechanisms within different visual areas.

Frequency tagging with fMRI. Frequency tagging has been used to study the effect of selective attention in a number of EEG studies (Andersen et al. 2008; Muller et al. 2006; Schoenfeld et al. 2007; Toffanin et al. 2009). To our knowledge, frequency tagging has not been used to study the effects of feature- or object-based attention with fMRI. However, frequency tagging is not new to fMRI; in traditional retinotopic mapping experiments, the spatial location of the stimulus is modulated in a circular fashion, both radially and tangentially, to infer the underlying retinotopic map from the phase of each voxel's response at the modulation frequency. The success of frequency tagging in retinotopic mapping is due to the large-scale cortical topography of receptive field locations in the visual cortex. This topography ensures that the profile of a voxel's response will oscillate at the frequency with which the retinal position of the stimulus is modulated.

The novel aspect of our work was to apply frequency tagging to analyze the responses to the features of color and direction of motion. We found that BOLD signals contained harmonic responses associated with the frequencies at which the color and direction of motion of the two surfaces were modulated. This finding suggests that some of the voxels contained an inhomogeneous distribution of feature-selective responses, for example, a weak preference for a particular color or direction of motion. The source of these inhomogeneities is not fully understood, but several possibilities have been proposed. Fine-scale anisotropies in the cortical organization of feature-selective neurons may give rise to the population response biases. Examples of such fine-scale anisotropies include the orientation-tuning columns and color-selective clusters in striate cortex (Hubel and Wiesel 1974; Xiao et al. 2007), the columnar structure of direction-selective neurons in MT (Albright 1984), and the color-selective clusters in extra-striate cortex (Conway et al. 2007). More recently, response biases at the voxel-based level were attributed to larger-scale biases in the topographic organization of the visual responses (Freeman et al. 2011). In addition, anisotropies in the underlying representation of simple features across voxels could result in a measurable harmonic response across a population of voxels (Freeman et al. 2011; Mannion et al. 2010; Op De Beeck 2010).

Several constraints must be satisfied to combine frequency tagging with fMRI measurements. The first constraint is related to the sluggish nature of the hemodynamic response. Based on a typical model of the hemodynamic impulse response function (Boynton et al. 1996), short temporal periods, like those previously used to incorporate frequency-tagging in EEG recordings (Morgan et al. 1996), would not be extractable from the BOLD signal. Consequently, because the BOLD signal cannot capture rapid modulations of the underlying neural activity, we chose relatively long temporal periods (on the order of tens of seconds).

A second constraint in our design was our use of circular feature spaces. To drive harmonic responses to a feature, it is important to be able to modulate that feature periodically. Color, direction motion, and orientation are all natural candidates because they can be readily represented and modulated in a circular fashion. It would be more difficult to use the frequency-tagging scheme to stimulus features that are not inherently circular, such as spatial or temporal frequency. We chose color and direction of motion because they are natural features of dot fields, which can be easily superimposed.

In addition, frequency tagging implicitly assumes a linear relationship between the stimulus and the evoked responses. Evidence suggests that the hemodynamic response is, to a first approximation, linearly related to the average population response over time (Boynton et al. 1996). This assumption underlies a large body of fMRI research that uses an estimate of the hemodynamic impulse response function to predict the BOLD response to a time-varying stimulus based on a general linear model (Heeger and Ress 2002). However, the assumption of linearity is not without challenge (Logothetis et al. 2001; Maier et al. 2008).

Specific to our stimulus, nonlinearities could also arise from neural populations tuned for specific color-motion conjunctions (Seymour et al. 2009). For example, if responses to the modulations of color and direction of motion interact multiplicatively, the harmonic responses would correspond to frequencies that are either higher or lower than those associated with the color and direction of motion in our stimulus. Consequently, our frequency analysis was insensitive to the output of neurons that combine color and motion nonlinearly.

Neural correlates of feature- and object-based attention. We measured fMRI responses while subjects tracked either the color or the direction of motion of one of two overlapping surfaces (dot fields) segregated by their unique conjunctions of color and motion. We designed the stimulus such that the four features (2 colors and 2 directions of motion) smoothly traversed a circular path through feature space with four unique temporal periods. This design enables us to use the corresponding harmonic responses to infer the effects of both feature- and object-based attention.

We found that the feature- and object-based attention effects were qualitatively similar (compare Fig. 3, *A* vs. *C*, and *B* vs. *D*). For example, A_{m1} (the amplitude of the harmonic associated with the motion of *surface 1*) was modulated in both the M1 and C1 conditions (Fig. 3, *A* and *C*). These results support the hypothesis that object-based attention modulates the sensory representation of all the features that comprise a task-relevant surface. Previous fMRI measurements found a correlate of object-based attention at the level of average BOLD signal across visual cortical areas (O'Craven et al. 1999). Our application of frequency tagging to fMRI data further shows that the mechanisms of object-based attention might operate at a finer level of feature-selective mechanisms, as suggested by electrophysiological measurements in nonhuman primates (Fallah et al. 2007; Katzner et al. 2009; Roelfsema et al. 1998; Wannig et al. 2007).

Moreover, our results reveal a correlate of object-based attention in V1. Although feature-based attention effects have

been reported in V1 (Saenz et al. 2002), object-based attention effects have only been reported in extrastriate visual areas, including MT+ (O'Craven et al. 1997; Katzner et al. 2009), the fusiform face area, and the parahippocampal place area (O'Craven et al. 1997). Our results extend previous work and suggest that the feedback signals mediated by object-based attention can target the representation of feature-specific mechanisms throughout the visual cortex, and as early as V1.

The attention indexes we used to quantify the effects of attention were based on comparing the hemodynamic responses to a feature under two different attentional states. This relative measure cannot differentiate between enhancement of responses to the attended feature and suppression of responses to the unattended feature (or a combination thereof). Distinguishing between these possibilities requires an estimate of the baseline response to the stimulus. We attempted to measure a baseline response to our stimulus in a separate fixation scan in which subjects performed a demanding fixation task to draw their attention away from either surface. Interestingly, with attention directed to fixation, the stimulus failed to produce a reliable response at any of the stimulus harmonics; the amplitude at the stimulus frequencies was indistinguishable from the surrounding noise frequencies. Therefore, we were unable to determine the nature of attentional modulations when attention was oriented toward one of the two surfaces.

Signal-to-noise ratio in our measurement. The average attention index for both the cued feature and the task-irrelevant feature on the cued surface was positive in all five regions of interest. However, our effect size was small. Several factors could have contributed to a weak overall harmonic response. First, it is thought that the superimposition of competing features within the receptive field of feature-selective neurons could reduce sensory-evoked responses (Desimone 1998; Moran and Desimone 1985; Reynolds et al. 1999). Such suppressive effects are likely to degrade the signal-to-noise ratio (SNR) of our measurements. A second contributing factor may be the nonlinearities in feature-selective responses. Our frequency-tagging approach is only able to extract signals that are driven by a linear combination of responses to individual features. Therefore, the inherent nonlinearities in sensory representations and the potential interactions between the four features in our stimuli may have further reduced the sensitivity of our measurements to feature- and object-based attentional modulations. Third, because of the sluggish nature of the hemodynamic response, it was necessary to modulate the color and direction of motion of the two surfaces using relatively long temporal periods. Such slow modulations could adapt central feature-selective mechanisms (Boynton and Finney 2003; Liu et al. 2007) and reduce the amplitude to the associated hemodynamic responses (Grill-Spector and Malach 2001).

The attention index for the motion harmonic was weaker in area MT+ than in earlier visual areas, a surprising result given that large effects of attention have been previously reported in area MT+ (O'Craven et al. 1997; Saenz et al. 2002; Serences and Boynton 2007; Tootell et al. 1998). This unexpected result might be related to a suppressive interaction between the two surfaces as shown by electrophysiological recordings in the macaque monkey (Treue et al. 2000), which is consistent with the weak harmonic responses our stimuli evoked in area MT+ (Fig. 7).

Across visual areas, the attention index was smaller for the color harmonics than for the motion harmonics. One possibility is that the distribution of color-selective neurons might be more homogeneous within a voxel than the distribution of motion-selective cells, effectively leading to a smaller color harmonic. Color has been successfully classified in a number of studies using fMRI (Brouwer and Heeger 2009; Kamitani and Tong 2006; Seymour et al. 2009). We used relatively large voxels (~3.5 mm isotropic) in our functional scans. The aforementioned pattern classification studies used 3-, 3-, and 2-mm isotropic voxels, respectively, in their functional scans. The volume of our voxels was therefore 1.7–5.9 times larger. This may have contributed to a more homogenous voxel response to color, thus reducing our overall SNR.

Another possibility might be related to the constraints of creating isoluminant stimuli, which limited the range of intensities we were able to use to modulate the red and green channels (to balance the luminance of the weaker blue channel). In addition, it is possible that direction of motion is inherently more effective than color in segregating transparent surfaces. If so, it is possible that even when subjects were asked to track the color of a surface, they still implicitly used the motion cue to improve their ability to segregate the two surfaces.

Behavioral strategy. We asked subjects to attend one feature (e.g., color) of one surface and found that attention modulated responses to the other feature of the same surface (e.g., motion). Although this result supports our hypothesis of spread of attention across features of a surface, it is important to ensure that the spread of attention was not inadvertently motivated by our task design or the choice of stimulus. To examine such potential confounding factors, we considered different behavioral strategies that were consistent with the behavioral and fMRI results.

First, we asked whether subjects could have used a behavioral strategy based on spatial attention to track the cued feature. Although unlikely, it is conceivable that subjects tracked the direction of motion indirectly using a spatial strategy; for example, subjects could have tracked a subset of dots that moved radially, away from the fixation point (i.e., orthogonal to the outer edge of the stimulus). In this strategy, the locus of attention revolves around the fixation point at the same stimulus frequency and leads to a harmonic response due to spatial (not feature) attention. However, this does not seem to be the case in our experiment because the phase maps associated with the attentional modulations did not resemble the phase maps derived from the retinotopy scans.

Second, we asked whether our experimental design inadvertently motivated the subjects to switch attention to the noncued feature, even though it was irrelevant. This also seems unlikely because subjects did not have to detect any event associated with the noncued feature. However, it is conceivable that subjects occasionally found it advantageous to switch attention to the noncued feature to facilitate tracking the target surface. This is particularly relevant during the windows of time when the attended feature in the target and distractor surfaces is similar (i.e., when the two surfaces "collide" in the feature space). For example, when motion is the relevant feature, during periods in which both surfaces move upward, subjects may find it advantageous to attend the color as a way to distinguish between the surfaces. We cannot rule out the

possibility that subjects did occasionally attend the noncued feature, but this possibility does not seem to explain two features of our data. First, because such collisions occurred infrequently, we would expect the effect of attention to the noncued feature to be significantly weaker than the effect of attention to the cued feature. In contrast, we found the effect of attention to the task-irrelevant feature to be large, sometimes larger than the effect of attention to the cued feature. Second, this strategy cannot explain the key observation that behavioral performance was equivalent regardless of whether one or both features were cued (Fig. 9).

We also considered various other behavioral strategies, but none could adequately explain our results. We therefore suggest that the modulation of responses to the task-irrelevant feature corresponds to an automatic spread of attention from the cued to the noncued feature of the same surface.

Perceptual consequences of object-based attention. Object-based attention allows subjects to divide attention to multiple features of an object with no additional cost. Our fMRI experiment showed that attention to one feature of a surface modulated the neural response to that feature as well as to the other feature of the same surface, but our behavioral paradigm did not directly test the consequences of divided attention. We therefore performed an additional psychophysical experiment to establish a more direct link between our fMRI measurements and the behavioral effects of object-based attention. Following previous work (Blaser et al. 2000; Bonnel and Prinzmetal 1998; Duncan 1984), we looked for a dual-task deficit when attention was divided between surfaces. In the single-task condition, subjects were required to detect changes in one of the surface features, and in the dual-task condition, they divided their attention across both features within one of the surfaces. We found that subjects' performance was comparable between the two conditions (Fig. 9), suggesting that capacity is not limited when attention is divided between features within a surface. This result is consistent with previous object-based divided attention experiments (Blaser et al. 2000; Bonnel and Prinzmetal 1998; Duncan 1984; Ernst et al. 2012). In conjunction with our fMRI measurements, these results suggest that the unfettered ability to attend to multiple features of a surface may be due to the simultaneous enhancement of sensory responses to the features of that surface early in the visual processing stream.

ACKNOWLEDGMENTS

GRANTS

DISCLOSURES

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

AUTHOR CONTRIBUTIONS

Z.R.E. and M.J. conception and design of research; Z.R.E. performed experiments; Z.R.E. and M.J. analyzed data; Z.R.E., G.M.B., and M.J. interpreted results of experiments; Z.R.E. and M.J. prepared figures; Z.R.E. drafted manuscript; Z.R.E., G.M.B., and M.J. edited and revised manuscript; Z.R.E., G.M.B., and M.J. approved final version of manuscript.

REFERENCES

Andersen SK, Hillyard SA, Muller MM. Attention facilitates multiple stimulus features in parallel in human visual cortex. *Curr Biol* 18: 1006–1009, 2008.

Anstis P, Cavanagh P. *A Minimum Motion Technique for Judging Equiluminance.* London: Academic, 1983.

Blaser E, Pylyshyn ZW, Holcombe AO. Tracking an object through feature space. *Nature* 408: 196–199, 2000.

Bonnel AM, Prinzmetal W. Dividing attention between the color and the shape of objects. *Percept Psychophys* 60: 113–124, 1998.

Boynton GM, Engel SA, Glover GH, Heeger DJ. Linear systems analysis of functional magnetic resonance imaging in human V1. *J Neurosci* 16: 4207–4221, 1996.

Boynton GM, Finney EM. Orientation-specific adaptation in human visual cortex. *J Neurosci* 23: 8781–8787, 2003.

Boynton GM. Imaging orientation selectivity: decoding conscious perception in V1. *Nat Neurosci* 8: 541–542, 2005a.

Boynton GM. Attention and visual perception. *Curr Opin Neurobiol* 15: 465–469, 2005b.

Brouwer GJ, Heeger DJ. Decoding and reconstructing color from responses in human visual cortex. *J Neurosci* 29: 13992–14003, 2009.

Chen AJW, Britton M, Turner GR, Vytlacil J, Thompson TW, D'Esposito M. Goal-directed attention alters the tuning of object-based representations in extrastriate cortex. *Front Hum Neurosci* 6: 187, 2012.

Desimone R. Visual attention mediated by biased competition in extrastriate visual cortex. *Philos Trans R Soc Lond B Biol Sci* 353: 1245–1255, 1998.

Duncan J. Selective attention and the organization of visual information. *J Exp Psychol* 113: 501–517, 1984.

Ernst ZR, Palmer J, Boynton GM. Dividing attention between two transparent motion surfaces results in a failure of selective attention. *J Vis* 12: 6, 2012.

Fallah M, Stoner GR, Reynolds JH. Stimulus-specific competitive selection in macaque extrastriate visual area V4. *Proc Natl Acad Sci USA* 104: 4165–4169, 2007.

Freeman J, Brouwer GJ, Heeger DJ, Merriam EP. Orientation decoding depends on maps, not columns. *J Neurosci* 31: 4792–4804, 2011.

Grill-Spector K, Malach R. fMR-adaptation: a tool for studying the functional properties of human cortical neurons. *Acta Psychol (Amst)* 107: 293–321, 2001.

Grill-Spector K, Sayres R. Object Recognition. *Curr Dir Psychol Sci* 17: 73–79, 2008.

Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293: 2425–2430, 2001.

Haynes JD, Rees G. Predicting the orientation of invisible stimuli from activity in human primary visual cortex. *Nat Neurosci* 8: 686–691, 2005.

Heeger DJ, Ress D. What does fMRI tell us about neuronal activity? *Nat Rev Neurosci* 3: 142–151, 2002.

Kamitani Y, Tong F. Decoding the visual and subjective contents of the human brain. *Nat Neurosci* 8: 679–685, 2005.

Kamitani Y, Tong F. Decoding seen and attended motion directions from activity in the human visual cortex. *Curr Biol* 16: 1096–1102, 2006.

Katzner S, Busse L, Treue S. Attention to the color of a moving stimulus modulates motion-signal processing in macaque area MT: evidence for a unified attentional system. *Front Syst Neurosci* 3: 12, 2009.

Liu T, Larsson J, Carrasco M. Feature-based attention modulates orientation-selective responses in human visual cortex. *Neuron* 55: 313–323, 2007.

Logothetis NK, Pauls J, Augath M, Trinath T, Oeltermann A. Neurophysiological investigation of the basis of the fMRI signal. *Nature* 412: 150–157, 2001.

Maier A, Wilke M, Aura C, Zhu C, Ye FQ, Leopold DA. Divergence of fMRI and neural signals in V1 during perceptual suppression in the awake monkey. *Nat Neurosci* 11: 1193–1200, 2008.

Mannion DJ, McDonald JS, Clifford CW. Orientation anisotropies in human visual cortex. *J Neurophysiol* 103: 3465–3471, 2010.

Martinez-Trujillo JC, Treue S. Feature-based attention increases the selectivity of population responses in primate visual cortex. *Curr Biol* 14: 744–751, 2004.

Moran J, Desimone R. Selective attention gates visual processing in the extrastriate cortex. *Science* 229: 782–784, 1985.

Morgan ST, Hansen JC, Hillyard SA. Selective attention to stimulus location modulates the steady-state visual evoked potential. *Proc Natl Acad Sci USA* 93: 4770–4774, 1996.

Muller MM, Andersen S, Trujillo NJ, Valdes-Sosa P, Malinowski P, Hillyard SA. Feature-selective attention enhances color signals in early visual areas of the human brain. *Proc Natl Acad Sci USA* 103: 14250–14254, 2006.

O'Craven KM, Downing PE, Kanwisher N. fMRI evidence for objects as the units of attentional selection. *Nature* 401: 584–587, 1999.

O'Craven KM, Rosen BR, Kwong KK, Treisman AM, Savoy RL. Voluntary attention modulates fMRI activity in human MT-MST. *Neuron* 18: 591–598, 1997.

O'Toole AJ, Jiang F, Abdi H, Haxby JV. Partially distributed representations of objects and faces in ventral temporal cortex. *J Cogn Neurosci* 17: 580–590, 2005.

Op De Beeck H. Against hyperacuity in brain reading: spatial smoothing does not hurt multivariate fMRI analyses? *Neuroimage* 49: 1943–1948, 2010.

Regan D. *Human Brain Electrophysiology: Evoked Potentials and Evoked Magnetic Fields in Science and Medicine.* New York: Elsevier, 1989.

Reynolds JH, Chelazzi L, Desimone R. Competitive mechanisms subserve attention in macaque areas V2 and V4. *J Neurosci* 19: 1736–1753, 1999.

Roelfsema PR, Lamme VAF, Spekreijse H. Object-based attention in the primary visual cortex of the macaque monkey. *Nature* 395: 376–381, 1998.

Saenz M, Buracas GT, Boynton GM. Global effects of feature-based attention in human visual cortex. *Nat Neurosci* 5: 631–632, 2002.

Schoenfeld MA, Hopf JM, Martinez A, Mai HM, Sattler C, Gasde A, Heinze HJ, Hillyard SA. Spatio-temporal analysis of feature-based attention. *Cereb Cortex* 17: 2468–2477, 2007.

Serences JT, Boynton GM. Feature-based attentional modulations in the absence of direct visual stimulation. *Neuron* 55: 301–312, 2007.

Seymour K, Clifford CW, Logothetis NK, Bartels A. The coding of color, motion, and their conjunction in the human visual cortex. *Curr Biol* 19: 177–183, 2009.

Toffanin P, de Jong R, Johnson A, Martens S. Using frequency tagging to quantify attentional deployment in a visual divided attention task. *Int J Psychophysiol* 72: 289–298, 2009.

Tootell RB, Reppas JB, Kwong KK, Malach R, Born RT, Brady TJ, Rosen BR, Belliveau JW. Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *J Neurosci* 15: 3215–3230, 1995.

Tootell RB, Hadjikhani N, Hall EK, Marrett S, Vanduffel W, Vaughan JT, Dale AM. The retinotopy of visual spatial attention. *Neuron* 21: 1409–1422, 1998.

Treue S, Hol K, Rauber HJ. Seeing multiple directions of motion-physiology and psychophysics. *Nat Neurosci* 3: 270–276, 2000.

Wannig A, Rodríguez V, Freiwald WA. Attention to surfaces modulates motion processing in extrastriate area MT. *Neuron* 54: 639–651, 2007.

Zeki S, Watson JD, Lueck CJ, Friston KJ, Kennard C, Frackowiak RS. A direct demonstration of functional specialization in human visual cortex. *J Neurosci* 11: 641–649, 1991.