

Attention to Stimulus Features Shifts Spectral Tuning of V4 Neurons during Natural Vision

Stephen V. David,^{1,3,5} Benjamin Y. Hayden,^{2,3,6} James A. Mazer,^{4,7} and Jack L. Gallant^{3,4,*}

¹Program in Bioengineering

²Department of Molecular and Cellular Biology

³Program in Neuroscience

⁴Department of Psychology

University of California, Berkeley, Berkeley, CA 94720-1650, USA

⁵Present address: 2202 A.V. Williams Building, University of Maryland, College Park, MD 20742, USA

⁶Present address: 353B Levine Science Building, Duke University, Durham, NC 27710, USA

⁷Present address: Department of Neurobiology, Yale School of Medicine, Box 208001, New Haven, CT 06520, USA

*Correspondence: gallant@berkeley.edu

DOI 10.1016/j.neuron.2008.07.001

SUMMARY

Previous neurophysiological studies suggest that attention can alter the baseline or gain of neurons in extrastriate visual areas but that it cannot change tuning. This suggests that neurons in visual cortex function as labeled lines whose meaning does not depend on task demands. To test this common assumption, we used a system identification approach to measure spatial frequency and orientation tuning in area V4 during two attentionally demanding visual search tasks, one that required fixation and one that allowed free viewing during search. We found that spatial attention modulates response baseline and gain but does not alter tuning, consistent with previous reports. In contrast, feature-based attention often shifts neuronal tuning. These tuning shifts are inconsistent with the labeled-line model and tend to enhance responses to stimulus features that distinguish the search target. Our data suggest that V4 neurons behave as matched filters that are dynamically tuned to optimize visual search.

INTRODUCTION

Cortical area V4 is an extrastriate visual area critical for form and shape perception (Gallant et al., 1993, 2000; Ogawa and Komatsu, 2004; Pasupathy and Connor, 1999; Schiller and Lee, 1991). Responses of V4 neurons are modulated by both spatial (Luck et al., 1997; Maunsell and Cook, 2002; McAdams and Maunsell, 1999; Moran and Desimone, 1985; Motter, 1993; Reynolds and Chelazzi, 2004) and feature-based attention (Hayden and Gallant, 2005; Mazer and Gallant, 2003; McAdams and Maunsell, 2000; Motter, 1994; Ogawa and Komatsu, 2004). Previous neurophysiological studies have reported that spatial attention directed into the receptive field (RF) of a V4 neuron can increase baseline firing rate, gain, or contrast sensitivity, but has little effect on feature selectivity (Luck et al., 1997; McAdams

and Maunsell, 1999; Ogawa and Komatsu, 2004; Reynolds et al., 2000; Williford and Maunsell, 2006). Studies of feature-based attention have been fewer in number. One important early study suggested that feature-based attention might change color selectivity, resulting in increased sensitivity to behaviorally relevant features (Motter, 1994). However, subsequent studies of feature-based attention in area V4 (McAdams and Maunsell, 2000) and MT (Martinez-Trujillo and Treue, 2004) have reported only changes in response gain. In this study, we used a new experimental and modeling approach to determine whether feature-based attention does, in fact, shift the visual tuning of V4 neurons.

Although neurophysiological evidence for tuning shifts in V4 is limited, a few theoretical and psychophysical studies have suggested that visual search might make use of a matched filter that shifts neuronal tuning toward the attended target (Carrasco et al., 2004; Compte and Wang, 2006; Lee et al., 1999; Lu and Doshier, 2004; Olshausen et al., 1993; Rao and Ballard, 1997; Tsotsos et al., 1995). Consistent with this idea, several reports have demonstrated that spatial attention can shift spatial receptive fields in both V4 and MT toward the attended location in the visual field (Connor et al., 1996, 1997; Tolias et al., 2001; Womelsdorf et al., 2006). For feature-based attention, the matched-filter hypothesis predicts that when attention is directed toward a spectral feature (e.g., a particular orientation or spatial frequency) neurons should shift their spectral tuning toward that feature. Evidence that attention alters tuning to visual features would challenge the classical, widely accepted hypothesis that neurons in visual cortex act as labeled lines with fixed tuning properties (i.e., the same optimal stimulus), regardless of attention state (Adrian and Matthews, 1927; Barlow, 1972; Marr, 1982).

Previous studies of attention in V4 may not have identified clear and compelling changes in visual tuning for two reasons. First, in order to maximize statistical power, most studies have used sparse stimulus sets of two to eight distinct, synthetic images that vary along one or two dimensions (Haenny et al., 1988; Luck et al., 1997; McAdams and Maunsell, 1999, 2000; Moran and Desimone, 1985; Motter, 1993; Reynolds et al., 2000). For these small stimulus sets, tuning changes are

detectable only when they fall directly along the dimension where the stimuli vary. Second, previous studies have focused primarily on the effects of spatial attention (Luck et al., 1997; McAdams and Maunsell, 1999; Reynolds et al., 2000). Space is represented topographically in extrastriate cortex while other features, such as orientation and spatial frequency, are not (Gattass et al., 1988). Thus, spatial attention may operate by a different mechanism and have different effects on tuning than attention to other features (Hayden and Gallant, 2005; Maunsell and Treue, 2006).

To investigate whether feature-based attention alters the tuning of V4 neurons, we performed two complementary experiments. These experiments used a spectrally rich set of natural images selected to fully span the likely tuning space encoded by V4 neurons, and they engaged both feature-based and spatial attention. The first, a match-to-sample task (MTS), allows simultaneous, independent manipulation of both spatial and feature-based attention in the absence of eye movements (Hayden and Gallant, 2005). The second, a free-viewing visual search task (FVVS), manipulates feature-based attention while allowing natural eye movements (Mazer and Gallant, 2003). Both tasks use natural images that broadly sample visual stimulus space and allow for the measurement of neuronal response properties under conditions approximating natural vision. We adopted a conservative analytical approach to identify tuning shifts; the statistical significance of shifts was assessed only after accounting for and removing any shifts in response baseline or gain. Other results from these experiments were reported previously (Hayden and Gallant, 2005; Mazer and Gallant, 2003).

RESULTS

To test the hypothesis that attention can alter tuning to visual features in area V4, we recorded responses from single neurons during an MTS task (Figure 1A; Hayden and Gallant, 2005) and a FVVS task (Figure 1B; Mazer and Gallant, 2003). In both tasks, feature-based attention was manipulated by specifying a natural image search target, and a large number of natural image distractors was presented in the receptive field of each neuron under each attention condition.

We characterized visual tuning by estimating the spectral receptive field (SRF; David et al., 2006) of each neuron from responses evoked by the distractors under different attention conditions. The SRF is a two-dimensional tuning profile that describes the joint orientation-spatial frequency tuning of a neuron (Mazer et al., 2002). The SRF provides a general second-order model of visual tuning, so it can be used to predict responses to arbitrary visual stimuli (David et al., 2006; Wu et al., 2006). Each SRF was estimated by normalized reverse correlation (Theunissen et al., 2001), a procedure that reliably characterizes tuning properties from responses to natural stimuli in both the visual (David and Gallant, 2005; David et al., 2006) and auditory systems (Woolley et al., 2005).

To determine how attention influences V4 SRFs during visual search, we evaluated three quantitative models of attentional modulation. (1) The *no modulation model* (Figure 2A) assumes that the estimated SRF is not affected by the state of attention. (2) The *baseline/gain modulation model* (Figure 2B) assumes

that attention modulates the baseline (i.e., firing rate in the absence of stimulation) or gain of each SRF but does not change spectral tuning (McAdams and Maunsell, 2000; Reynolds et al., 2000). Both the no modulation and baseline/gain modulation models are consistent with the labeled-line hypothesis. However, if the baseline/gain model is correct, then SRFs whose baseline and overall gain are fit separately for each attention condition should predict neuronal responses better than the no modulation model. (3) The *tuning modulation model* (Figure 2C) assumes that attention can modulate spectral tuning (i.e., the shape of the SRF), thereby changing a neuron's preferred stimulus. This model is inconsistent with the labeled-line hypothesis. If the tuning modulation model is correct, then SRFs estimated separately for each attention condition should predict neuronal responses better than either of the other two models. In our approach, these three models are designed hierarchically; each one is fit successively to the residual of the previous one (see *Experimental Procedures*). Thus, the tuning modulation model accounts only for changes in neuronal responses that cannot be attributed to baseline or gain modulation.

Feature-Based Attention Shifts Spectral Tuning in V4 during Match-to-Sample

To determine whether feature-based attention modulates spectral tuning during MTS, we collapsed the data over both conditions of spatial attention (into and away from the receptive field). The three models described above were then fit to the collapsed data from every neuron in the sample ($n = 105$). The performance of each model was evaluated using a strict crossvalidation procedure. This procedure precluded the possibility of overfitting to noise and allowed for unbiased comparison of models with different numbers of parameters (David and Gallant, 2005). Figure 3 shows SRFs and model fits for one V4 neuron whose tuning is modulated by feature-based attention during MTS. The SRF estimated for the no modulation model shows that this neuron is tuned for a narrow range of orientations and a wide range of spatial frequencies (75° – 90° and 5–10 cyc/RF). We quantified the performance of the no modulation model by measuring its prediction accuracy relative to the predictive power of the most comprehensive model tested (i.e., the tuning modulation model, see below). For this neuron, the no modulation SRF accounts for 77% of the total predicted response variance.

The effects of feature-based attention on response baseline and gain are visualized by plotting distractor responses predicted by the no modulation model against responses actually observed in each attention condition (Figure 3B). Because each point is averaged over only two stimulus presentations, there is substantial scatter due to neuronal response variability. When the data are binned across stimuli predicted to give similar responses (solid lines), the differences between attention conditions become more apparent. The change in slope indicates that this neuron shows greater gain when feature-based attention is directed to the *preferred target* (T1, i.e., the target that elicits the stronger visual response, averaged across attention conditions) than when it is directed to the *nonpreferred target* (T2). The baseline/gain modulation model accounts for an additional 8% of the total predicted response variance for this neuron,

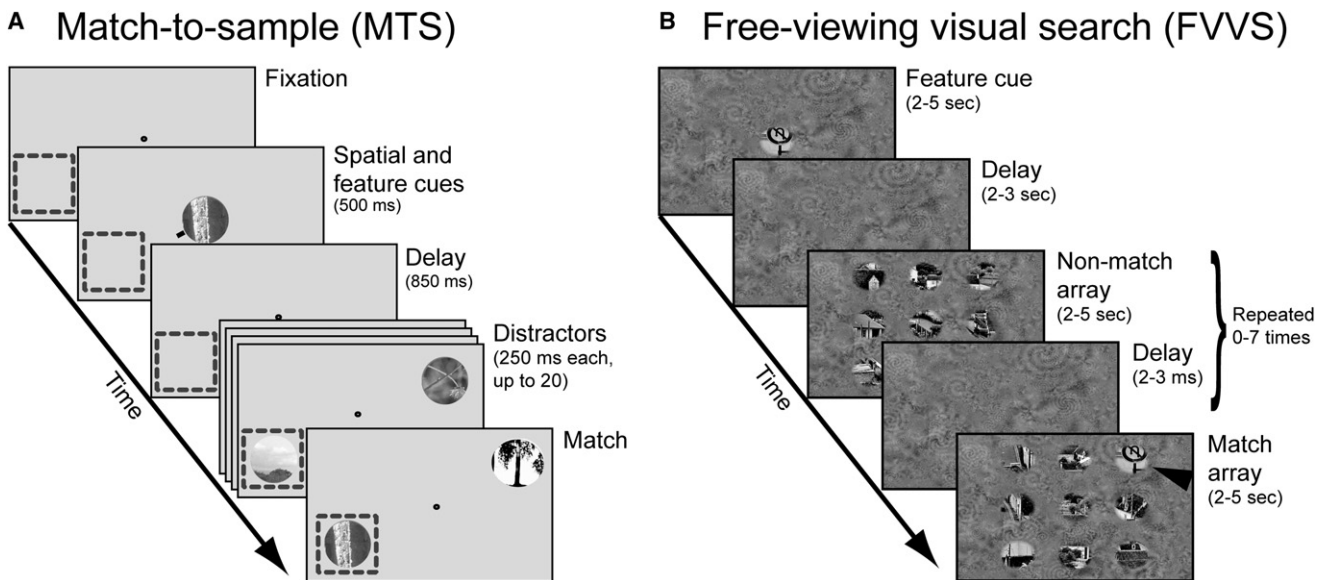


Figure 1. Behavioral Tasks

(A) Match-to-sample (MTS) task. Fixation was maintained at the center of the screen while a random, rapid sequence of natural images was presented. Spatial and feature-based attention were controlled independently on each trial: a spatial cue directed attention toward the receptive field of a V4 neuron (dashed box, not shown during the experiment) or toward the opposite hemifield; an image cue indicated the target. The task required a response when the cued target image appeared at the cued spatial location.

(B) Free-viewing visual search (FVVS) task. Eye movements were permitted while a random sequence of natural image arrays was shown. Feature-based attention was controlled with an image cue before array onset. (During the actual experiment, both the image patches and the background texture were shown at the same RMS contrast; the contrast of the background has been reduced here for illustrative purposes.) The task required a response whenever the sample image appeared in one of the image arrays (black arrowhead, not shown during the experiment).

a significant increase over the no modulation model ($p < 0.05$, jackknifed t test).

The tuning modulation model was fit by estimating SRFs separately for trials when T1 was the target and when T2 was the target (Figure 3C). This model also shows a significant increase in predictive power over the baseline/gain modulation model, accounting for the remaining 15% of total predicted response variance ($p < 0.05$, jackknifed t test). The shift in spectral tuning between attention conditions suggests a strategy that facilitates target detection: during T1 trials, spatial frequency tuning is sharper and higher than during T2 trials. Inspection of the power spectra of the target images (Figure 3C, bottom row) reveals that T1 has more power than T2 at high spatial frequencies. Thus, feature-based attention appears to shift spectral tuning to more closely match the spectral properties of the target. To test for this effect quantitatively, we computed a *target similarity index* (TSI) from the SRFs measured under each condition of feature-based attention (see Experimental Procedures). TSI values significantly greater than zero indicate that the SRF shifts toward the power spectrum of the search target. (More specifically, when comparing SRFs estimated separately from T1 and T2 trials, a positive TSI indicates that the SRF estimated from T1 trials is more like the spectrum of T1 and/or that the SRF estimated from T2 trials is more like the spectrum of T2.) An extreme value of 1 indicates a perfect match between the SRF and target spectra in both conditions of feature-based attention (i.e., a perfect matched filter). A tuning shift away from the target will produce a negative TSI, and a shift in any direction orthogonal to the tar-

get axis will produce a TSI of zero. For this neuron, the TSI of 0.17 is significantly greater than zero ($p < 0.05$, jackknifed t test), indicating a significant shift toward the search target.

Figure 4 shows another V4 neuron whose spectral tuning is also altered by feature-based attention. The no modulation SRF shows that this neuron is broadly tuned for orientation and narrowly tuned for spatial frequency, accounting for 30% of total predicted response variance (Figure 4A). On T1 trials response gain increases (baseline/gain modulation model, 9% of total predicted response variance, Figure 4B, $p < 0.05$) and spectral tuning shifts toward higher spatial frequencies, as compared to T2 trials (tuning modulation model, 61% of total predicted response variance, Figure 4C, $p < 0.05$). As in the previous example, the observed SRF changes are correlated with the spectral differences between T1 and T2; T1 has more power at low spatial frequencies than does T2. This neuron enhances tuning at low frequencies on trials when attention is directed toward T1, but its tuning at higher spatial frequencies remains unchanged (TSI: 0.04, $p < 0.05$, jackknifed t test).

Feature-based attention modulates baseline and gain in about half of the neurons in our sample (55/105, 52%, $p < 0.05$, jackknifed t test, Figure 5A), a finding consistent with previous studies (McAdams and Maunsell, 2000; Motter, 1994) and consistent with the labeled-line hypothesis. However, feature-based attention also shifts the spectral tuning of nearly one-third of the neurons, a finding that is inconsistent with the labeled-line hypothesis (31/105, 30%, $p < 0.05$, jackknifed t test). In order to explain the observed effects of feature-based attention, a complete

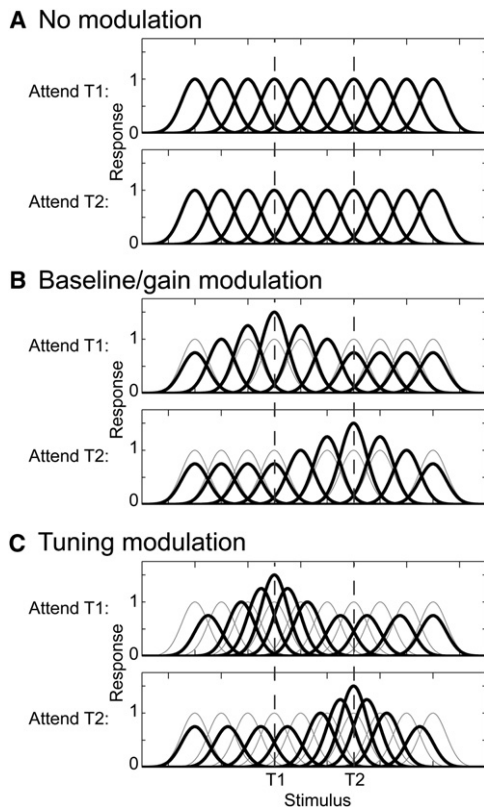


Figure 2. Alternative Models of Modulation by Attention

(A) The *no modulation model* assumes that neuronal responses are not affected by attention. The set of Gaussian curves represents the tuning of a population of neurons in an arbitrary stimulus space (horizontal axis). T1 and T2 represent two points where attention can be directed in the stimulus space. Tuning curves are the same whether attention is directed to T1 (top panel) or T2 (bottom panel).

(B) The *baseline/gain modulation model* assumes that attention modulates mean responses or response gain but does not change tuning. This model is consistent with a labeled-line code.

(C) The *tuning modulation model* assumes that attention can shift (or reshape) tuning curves of individual neurons. This model is inconsistent with a labeled-line code. If neurons behave as matched filters, then tuning should shift toward the attended target.

model of attention must account for changes in the shape of the tuning profile in addition to changes in response baseline and gain.

Figure 5B compares the contribution of different attention-dependent models to the predictive power of SRFs for all neurons significantly modulated by attention ($n = 87/105$). Modulation of response baseline and gain by feature-based attention (baseline/gain modulation model) accounts for an average of 19% of the total predicted response variance. Modulation of spectral tuning by feature-based attention (tuning modulation model) accounts for an additional 14% of the predicted response variance. (The remaining bulk of the predicted response variance is accounted for by the no modulation model.)

Note that the tuning shifts that we report here are conservative. In our crossvalidation procedure, each model is fit using one data set and then evaluated in terms of its predictive power

in an entirely separate validation data set. Therefore, our results represent a lower bound on the true proportion of neurons that show tuning shifts due to feature-based attention and on the magnitude of the shifts themselves.

To determine whether spatial attention also modulates spectral tuning during MTS, we collapsed data for each neuron over both feature-based attention conditions (T1 and T2) and fit the three models described above. For the neuron shown in Figure 4, the no modulation model accounts for 23% of total predicted response variance (Figure 4A). Spatial attention significantly modulates both response baseline and gain (77% of response variance, Figure 4D, $p < 0.05$), but does not alter the shape of the SRF ($p > 0.2$). Therefore, the SRFs estimated in separate spatial attention conditions (Figure 4E) show no discernible tuning shifts. Because feature-based attention shifts the spectral tuning of this neuron but spatial attention does not, this example demonstrates that feature-based and spatial attention can have different effects on tuning in the same neuron.

Spectral tuning is shifted by spatial attention during MTS in less than 15% of the neurons in our sample (12/105, 12%, $p < 0.05$, jackknifed t test, Figure 5A). This fraction is significantly lower than observed for feature-based attention ($p < 0.001$, randomized paired t test). In contrast, nearly three-fourths of the neurons (73/105) show significant modulation of baseline and/or gain by spatial attention ($p < 0.05$, jackknifed t test), significantly more than for feature-based attention ($p < 0.001$, randomized paired t test). These differences are also reflected in the fraction of response variance accounted for by the baseline/gain and tuning modulation models (Figure 5B). For spatial attention, the baseline/gain modulation model accounts for an average of 38% of total explained variance, and the tuning modulation model accounts for only 7%. Thus, spatial attention imposes large modulations on the responses of V4 neurons, but, unlike feature-based attention, its effects are restricted to response baseline and gain.

Feature-Based Attention Shifts Spectral Tuning toward Target Features

The examples presented above suggest that feature-based attention shifts spectral tuning toward the spectrum of the attended search target. If true, such shifts would suggest that area V4 acts as a matched filter that enhances the representation of task-relevant information and reduces the representation of task-irrelevant channels (Figure 2C; Compte and Wang, 2006; Martinez-Trujillo and Treue, 2004; Rao and Ballard, 1997; Tsotsos et al., 1995). To test this hypothesis, we computed the target similarity index (TSI) for each neuron that showed significant spectral tuning modulation during MTS. SRFs estimated using the tuning modulation model for each of the 31 neurons showing significant tuning modulation appear in Figure S1 available online. About 50% of neurons (16/31) have TSIs significantly greater than zero ($p < 0.05$, jackknifed t test), while only one TSI is significantly less than zero ($p < 0.05$, jackknifed t test). The population average TSI of 0.13 is also significantly greater than zero ($p < 0.01$, jackknifed t test, Figure 6). Thus, V4 neurons tend to shift their tuning toward the attended feature, as predicted by the matched filter hypothesis.

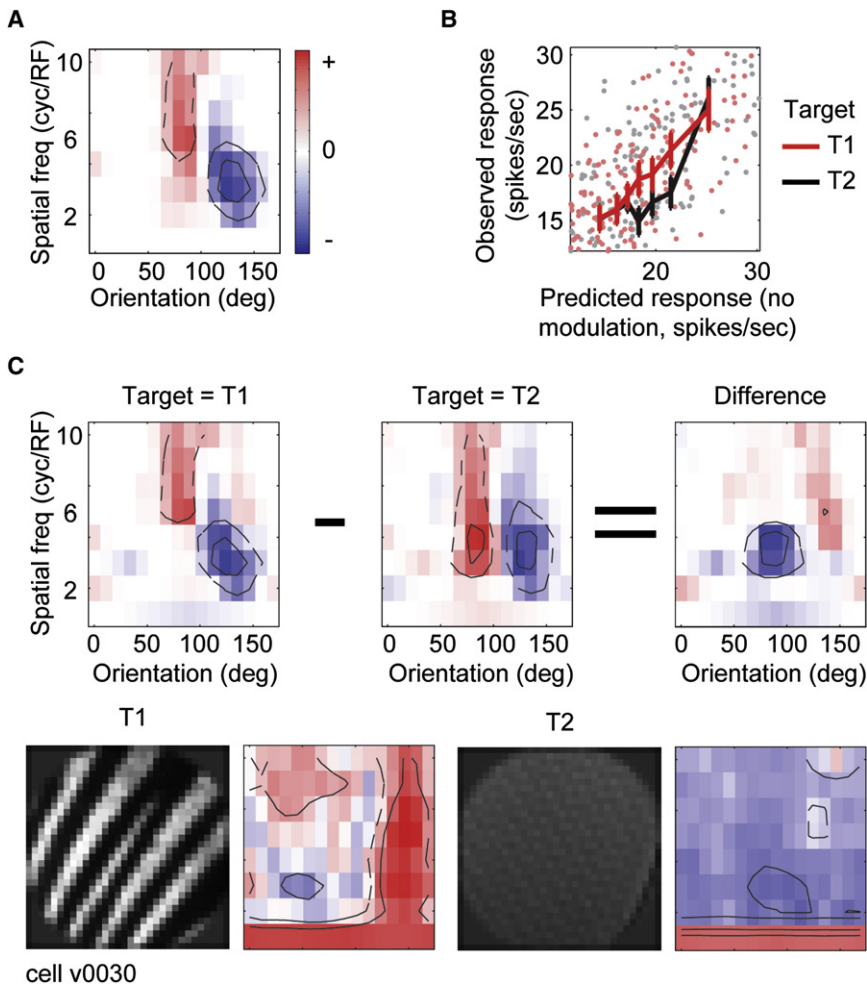


Figure 3. Representative V4 Neuron in which Spectral Tuning Is Modulated by Feature-Based Attention during MTS

(A) Spectral receptive field (SRF) estimated after averaging over all attention conditions (the no modulation model). Red regions indicate excitatory orientation and spatial frequency channels, and blue regions indicate suppressive channels. Contours enclose channels whose amplitude is one/two standard deviations above/below zero. This neuron is sharply tuned for orientation and broadly tuned for spatial frequency (75°–90° and 5–10 cyc/RF).

(B) To test for baseline/gain modulation, responses predicted by the no modulation SRF from panel (A) (horizontal axis) are plotted against observed responses (vertical axis) on trials when the target was T1 (red) or T2 (black). Solid lines show binned responses and error bars indicate one standard error of the mean. Responses on T1 trials are significantly greater than on T2 trials (jackknifed t test, $p < 0.05$).

(C) To test for tuning modulation, SRFs are estimated independently using data from only T1 (left) or T2 (center) trials (difference at right). The excitatory tuning channels shift to lower spatial frequencies on T2 trials (jackknifed t test, $p < 0.05$), indicating that feature-based attention modulates spectral tuning in this neuron. The four bottom panels illustrate targets T1 and T2 and their respective Fourier power spectra.

Data were analyzed by estimating SRFs for each neuron using the same three models used to evaluate the MTS data.

Figure 7 illustrates one V4 neuron whose spectral tuning depended on the

The examples shown in Figures 3 and 4 suggest that, during MTS, tuning shifts occur primarily along the spatial frequency dimension rather than the orientation dimension. However, the specific pattern of tuning modulation across area V4 neurons is diverse (see Figure S1), perhaps due to the substantial diversity of estimated SRFs (David et al., 2006). To determine whether matched filter effects occur mainly along the spatial frequency tuning dimension, we computed the TSI for each neuron with significant tuning modulation after collapsing the SRFs along the spatial frequency or orientation axis. When measured separately for these dimensions, the average TSI for spatial frequency tuning is 0.12 and the average for orientation tuning is 0.06. Both means are significantly greater than zero ($p < 0.01$, jackknifed t test). Few individual neurons show significant shifts, which is likely due to the limited signal-to-noise level available for this more fine-grained analysis.

Feature-Based Attention Modulates Spectral Tuning in V4 during Free-Viewing Visual Search

To determine whether feature-based attention modulates visual tuning under more naturalistic conditions, we analyzed a separate data set acquired in an FVVS task in which voluntary eye movements were permitted (Figure 1B; Mazer and Gallant, 2003).

search target during FVVS. The SRF estimated using the no modulation model shows that this neuron is tuned to vertical orientations and low spatial frequencies (92% of total predicted response variance, Figure 7A). Feature-based attention has no significant effect on either response baseline or gain (0% of total predicted response variance, Figure 7B). However, spatial frequency tuning on T1 trials is higher than on T2 trials (8% of total predicted response variance, Figure 7C, $p < 0.05$). Inspection of the power spectra of the target images (Figure 7C, bottom row) reveals that T1 has more power than T2 at vertical orientations and high spatial frequencies (TSI = 0.05, $p < 0.05$, jackknifed t test). Thus, feature-based attention shifts the tuning of this neuron to more closely match the spectral properties of the target during FVVS.

About 45% (39/87) of the V4 neurons in our sample show significant baseline/gain modulation during FVVS, and 25% (22/87) show significant spectral tuning shifts ($p < 0.05$, jackknifed t test, Figure 5A). The frequency of occurrence for both types of modulation under these more natural conditions is not significantly different than for the feature-based attention effects we observe during MTS (jackknifed t test). We also measured the fraction of response variance accounted for by each of the three attentional modulation models (Figure 5B). For the FVVS

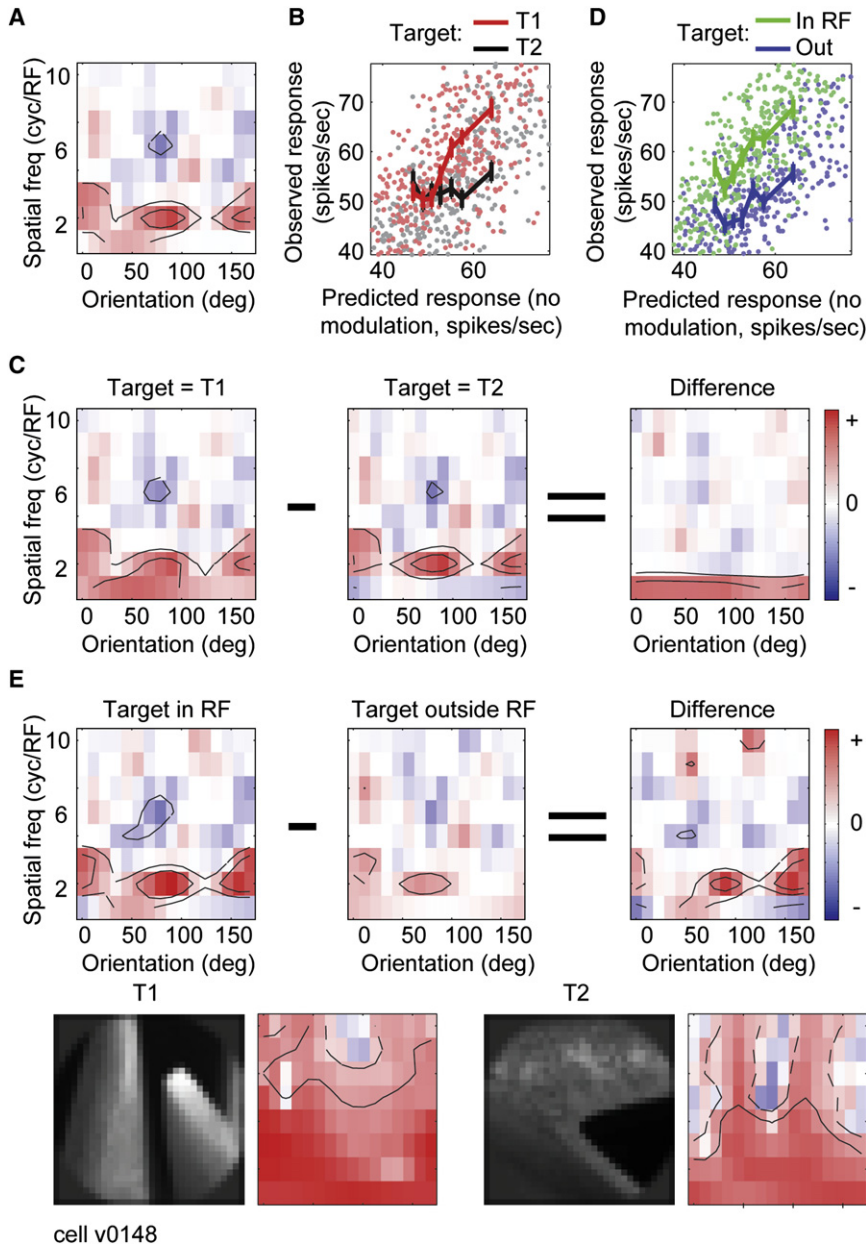


Figure 4. Representative V4 Neuron in which Spectral Tuning Is Modulated by Feature-Based Attention and Response Baseline/Gain Is Modulated by Both Feature-Based and Spatial Attention during MTS
Data are plotted as in Figure 3.

(A) SRF estimated after averaging over all attention conditions (the no modulation model). This neuron is tuned to a broad range of orientations at low spatial frequencies (0°–180° and 1–3 cyc/RF).

(B) Responses on T1 trials are larger than on T2 trials, indicating a significant modulation of baseline/gain by feature-based attention (jackknifed t test, $p < 0.05$). Error bars indicate one standard error of the mean value in each bin.

(C) Feature-based attention modulates spectral tuning in this neuron. The low spatial frequency tuning that appears in the SRF during T1 trials is absent during T2 trials, reflecting a shift in tuning toward higher spatial frequencies (jackknifed t test, $p < 0.05$).

(D) Responses when spatial attention is directed into the receptive field are stronger than when it is directed away, indicating a significant modulation of baseline/gain by spatial attention ($p < 0.05$). (E) Spatial attention does not modulate spectral tuning in this neuron. Aside from the global change in gain also observed in (D), the SRF estimated using trials when spatial attention is directed into the receptive field are not significantly different from those obtained when it is directed away, and not different from the SRF estimated under the no modulation model shown in panel (A).

conditions, and these differences were no larger for neurons that showed significant tuning modulation than for those that did not show tuning modulation.

To determine whether the shifts in spectral tuning observed during FVVS are also compatible with a matched filter, we measured TSI for the 22 neurons in our sample that showed significant shifts in spectral tuning. In contrast to what we observed during MTS, only four neurons had TSIs significantly greater than zero ($p < 0.05$, jackknifed t test), and none

data, the no modulation model accounts for an average of 78% of the total predicted response variance. The baseline/gain modulation model accounts for an average of 13%, and the tuning modulation model accounts for the remaining 9%. These figures are also not significantly different from those observed during MTS.

Eye movements were not controlled during FVVS. Thus, it is theoretically possible that the modulation observed during FVVS might reflect eye-movement-related differences across search conditions. To test this possibility, we compared mean fixation duration and saccade length during different attention conditions (Figure S2). We rarely observed significant differences in the pattern of eye movements between attention

were significantly less the zero. The average TSI of 0.007 for the FVVS data was not significantly greater than zero. This difference between the FVVS and MTS data could reflect differences in the effects of attentional mechanisms between tasks, but it is more likely that it simply reflects the relatively lower signal-to-noise level in FVVS SRF estimates (because of uncertainty in eye calibration [Mazer and Gallant, 2003]).

Feature-Based and Spatial Attention Are Additive and Independent

Previous neurophysiological (Hayden and Gallant, 2005; McAdams and Maunsell, 2000; Motter, 1994; Treue and Martinez-Trujillo, 1999) and psychophysical (Rossi and Paradiso, 1995;

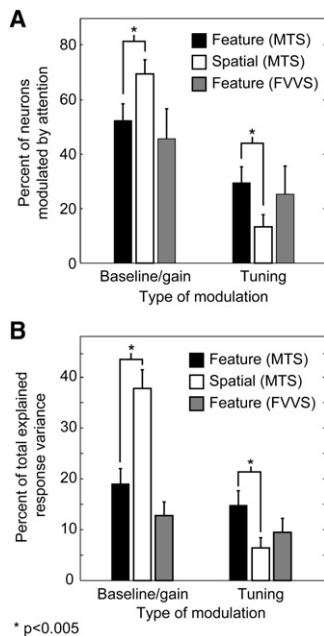


Figure 5. Frequency and Magnitude of Attention Effects

(A) Fraction of neurons modulated by attention. Error bars indicate one standard error. During MTS, significantly more neurons show baseline/gain modulation by spatial attention (white bars) than by feature-based attention (black bars, $p < 0.005$, randomized paired t test). Conversely, significantly more neurons show tuning modulation by feature-based attention than by spatial attention ($p < 0.005$, randomized paired t test). The number of neurons undergoing significant baseline/gain and tuning modulation by feature-based attention during FVVS (gray bars) is not significantly different from MTS (jackknifed t test).

(B) Average percent of total response variance explained by the baseline/gain modulation and tuning modulation models during MTS and FVVS tasks, for those neurons that show any significant effect of attention (MTS: 84/105; FVVS: 55/87). The remaining portion of response variance is explained by the no modulation model. Error bars indicate one standard error. Across the set of neurons, modulation of baseline and/or gain by spatial attention during MTS (white bars) accounts for a significantly greater portion of response variance than feature-based attention (black bars, $p < 0.005$, randomized paired t test). Tuning modulation by feature-based attention during MTS accounts for a significantly greater portion of response variance than tuning modulation by spatial attention ($p < 0.005$, randomized paired t test). The effects of feature-based attention during FVVS (gray bars) are not significantly different than during MTS (jackknifed t test).

Saenz et al., 2003) studies have reported that feature-based attention is deployed simultaneously throughout the visual field, independent of the locus of spatial attention. These reports suggest that the top-down influences of feature-based and spatial attention arise from separate networks that feed back into V4. If the two forms of attention are implemented by separate networks, then their effects on neuronal responses should be additive and independent.

During MTS, we find that both feature-based and spatial attention effects tend to co-occur in the same V4 neurons: neurons that are modulated by either feature-based or spatial attention are also modulated by the other form of attention significantly more often than would be expected if these two forms of attention operated on random subsets of neurons (Figure 8A,

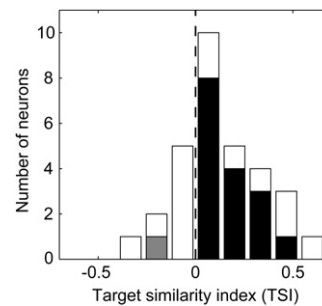


Figure 6. Evidence for a Matched Filter in V4

Histogram plots the tuning shift index (TSI) for neurons that show significant tuning modulation by feature-based attention during MTS ($n = 31/105$). Index values greater than zero indicate that spectral tuning shifts to match the spectrum of target images under different feature-based attention conditions. A value of 1 indicates a perfect match between SRF and target spectrum in both attention conditions. Black bars indicate TSIs significantly greater than 0 (16/31 neurons, $p < 0.05$, jackknifed t test), and gray bars indicate TSIs significantly less than 0 (1/31 neurons, $p < 0.05$). The mean TSI of 0.13 is significantly greater than zero ($p < 0.01$, jackknifed t test), indicating that tuning tends to shift to match the spectrum of the target. This increase in TSI is what would be expected if feature-based attention in area V4 implemented a matched filter.

$p < 0.001$). This suggests that some neurons, perhaps those more highly connected to central areas, are more often influenced by top-down processes, but these effects could still be additive and independent.

To test for dependency between feature-based and spatial attention directly, we measured the effect of feature-based attention separately when spatial attention was either directed toward or away from the receptive field. In neurons with high signal-to-noise ratios, tuning modulation by feature-based attention is similar when the data are analyzed separately for the two spatial attention conditions (Figure S3). The number of neurons in our sample that show significant modulation of baseline/gain and spectral tuning by feature-based attention is not significantly different when spatial attention is directed toward or away from the receptive field (Figure 8B). Therefore, feature-based and spatial attention tend to affect the same neurons, but their effects are largely additive and independent.

If feature-based and spatial attention operate through different feedback networks, they may also have different effects on response baseline and gain. To evaluate baseline/gain effects in more detail, we calculated the average visual response to each distractor during MTS (i.e., the average response collapsed across all conditions of feature-based and spatial attention) and plotted the change in response due to either feature-based or spatial attention as a function of the visual response (Figure 8C). For feature-based attention, we compared distractor responses on T1 (preferred) trials to T2 (nonpreferred) trials. Responses are generally enhanced when attention is directed to the preferred feature, but the enhancement is not uniform across stimuli. Responses to distractors that elicit larger responses show larger modulation, while responses to distractors that elicit weak responses are not typically modulated (red curve, Figure 8C). In contrast, spatial attention modulates responses to all distractors, including those that elicit only small responses

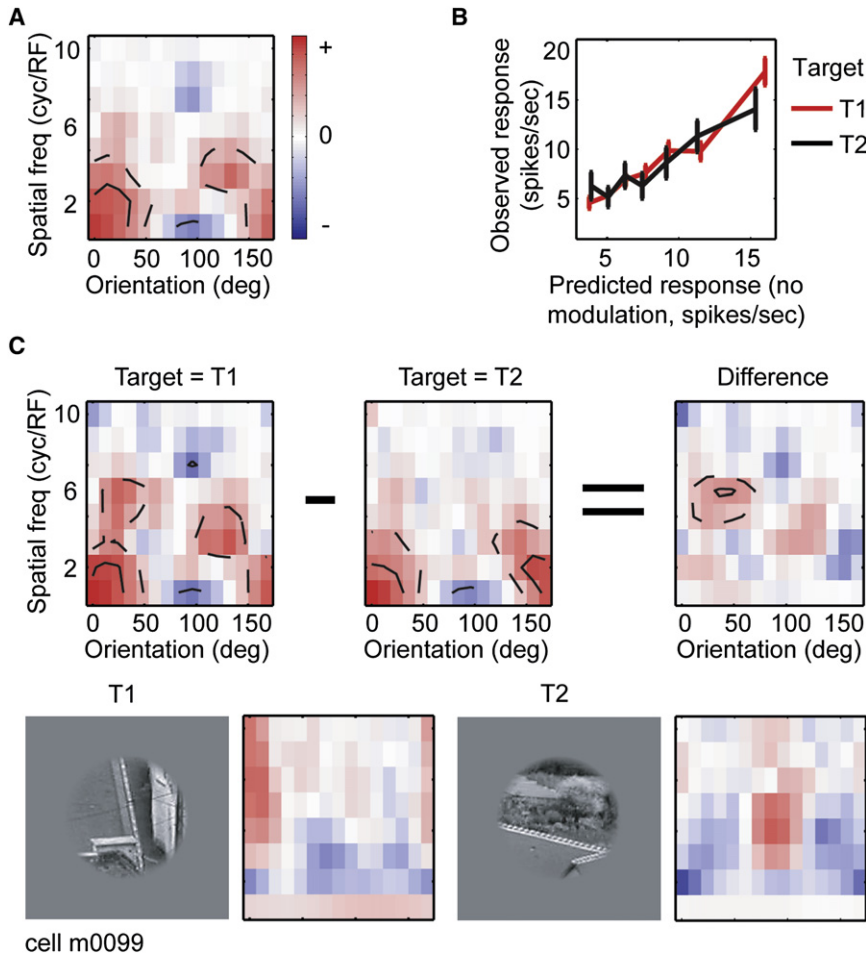


Figure 7. Representative V4 Neuron in which Spectral Tuning Is Modulated by Feature-Based Attention during FVVS

Data are plotted as in Figure 3.

(A) SRF estimated after averaging over both targets (the no modulation model). This neuron is tuned to vertical orientations and low spatial frequencies (170°–30° and 1–4 cyc/RF).

(B) Responses on T1 and T2 trials are equal, indicating that baseline and gain are not modulated by feature-based attention. Error bars indicate one standard error.

(C) Feature-based attention modulates spectral tuning in this neuron. The excitatory tuning channels are shifted significantly to lower spatial frequencies on T2 trials than on T1 trials (jackknifed t test, $p < 0.05$).

Comparison with Previous Studies of Attention-Mediated Tuning Shifts

Our results are unprecedented in their report of shifts in spectral tuning in V4 neurons. Two previous reports have suggested changes in either orientation (Haenny and Schiller, 1988) or color tuning (Motter, 1994) due to attention; however, there has been some question about whether the observed effects reflected a true change in feature tuning. The experiments in this study were designed specifically to control for the factors that limited interpretation of the earlier studies. In the first study, Haenny

and Schiller (1988) reported that orientation tuning bandwidth could change with spatial attention. A later study suggested that these results could also be explained by a change in response gain (McAdams and Maunsell, 1999). Therefore, in this study, we first measured attention-dependent changes in response baseline and gain and then measured changes in tuning that could not be explained by global baseline or gain modulation.

Motter (1994) reported that color selectivity in V4 changed as attention was shifted toward different colors. However, the design of the experiment was such that subjects could have used spatial attention in addition to feature-based attention to perform the task. As a result, it was theoretically possible that gain changes due to spatial attention might have caused apparent shifts in color selectivity. To avoid this potential confound, our study controlled spatial attention, varying both spatial and feature-based attention independently.

Subjects searching for a natural image might attend to the collection of features that together compose the image. Changes in spatial frequency tuning could be an effect of object-based attention (Fallah et al., 2007) or shrinking of the RF around a small feature in the target (Moran and Desimone, 1985; Motter, 1993). Despite the possibility that subjects could employ a range of strategies in the task, we still observe

(green curve, Figure 8C). This pattern suggests that feature-based attention modulates response gain but not response baseline while spatial attention modulates both baseline and gain (McAdams and Maunsell, 1999; Reynolds et al., 2000). This difference provides further support for the idea that feature-based and spatial attention are implemented in distinct neural circuits.

DISCUSSION

This study provides unambiguous evidence that feature-based attention can alter the spectral tuning of V4 neurons during natural vision. As reported in previous studies (Mazer and Gallant, 2003; McAdams and Maunsell, 2000), we observed changes in response baseline and gain due to feature-based attention. In addition, we found that many neurons can actually change their spectral tuning as attention is directed toward different features, and these changes cannot be explained by the baseline/gain modulation model. In neurons whose tuning is modulated by feature-based attention, tuning often shifts to more closely match the attended spectral feature (i.e., the search target). Our results suggest that the neural representation of shape in extrastriate visual cortex is dynamic and context dependent (Gilbert and Sigman, 2007).

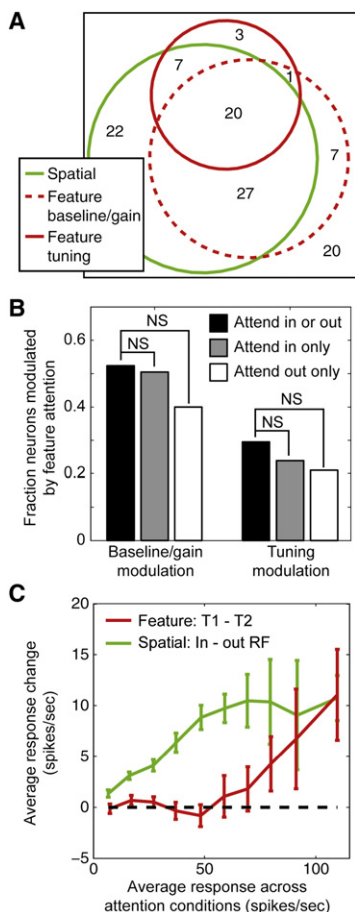


Figure 8. Effects of Feature-Based and Spatial Attention Are Additive and Independent

(A) Venn diagram summarizes the overlap of neurons with significant modulation by feature-based attention (baseline/gain and tuning modulation shown separately) and by spatial attention. A large proportion of the neurons are modulated by both forms of attention (83%), which is significantly greater than chance (random percent overlap: 72%, jackknifed t test, $p < 0.001$). Of neurons that are modulated significantly by feature-based attention, a significant proportion show changes both in baseline/gain responses and in spectral tuning (68%, random overlap: 52%, jackknifed t test, $p < 0.05$).

(B) Test for dependency between the effects of feature-based and spatial attention on response properties. At left, bars show the fraction of V4 neurons that show significant baseline/gain modulation for data averaged over both conditions of spatial attention and for each spatial attention condition separately. The number of significantly modulated cells ($p < 0.05$, jackknifed t test) is largest when data are averaged across spatial attention conditions (black bar), indicating that the effects of feature-based attention are the same in both spatial conditions and that averaging across spatial conditions improves the signal-to-noise ratio of recorded responses. The number of modulated neurons is slightly larger when spatial attention is directed into the receptive field (gray bar) than away from the receptive field (white bar), but the difference is not significant (jackknifed t test). Similarly, the fraction of neurons showing tuning modulation by feature-based attention was not significantly different between spatial attention conditions (bars at right).

(C) Comparison of baseline and gain modulation by feature-based and spatial attention. Red curve gives the difference in response to distractors in preferred (T1) versus nonpreferred (T2) feature conditions during MTS, as a function of the response averaged across attention conditions ($n = 87$ neurons that undergo any significant modulation by attention). Error bars indicate one stan-

dard error of the mean in each bin. Responses are enhanced during T1 trials only for stimuli that elicit strong responses under both feature-based attention conditions. Green curve gives the difference in response to distractors when spatial attention is directed into and away from the receptive field, as a function of the average response ($n = 87$). Responses are enhanced when attention is directed into the receptive field, regardless of the effectiveness of the stimuli.

significant tuning shifts in addition to changes in response baseline and gain. Our observation that spectral tuning shifts occur regardless of whether spatial attention is directed into or away from the receptive field suggests that at least some of these shifts are mediated by a global feature-based mechanism. It is important to note that the majority of neurophysiological studies of attention in area V4 have focused on the effects of spatial attention rather than feature-based attention (Desimone and Duncan, 1995; McAdams and Maunsell, 1999; Reynolds et al., 2000; Williford and Maunsell, 2006). The effects of spatial attention reported here are generally consistent with previous studies, confirming that spatial attention modulates response baseline and gain but has little effect on spectral tuning. Two spatial attention studies did report shifts in the spatial envelope of V4 receptive fields (Connor et al., 1997; Tolias et al., 2001), a spatial effect analogous to the matched filter effects reported here.

One important limitation of previous studies is that most only measured responses to a relatively small number of distinct stimuli (Haenny et al., 1988; Luck et al., 1997; McAdams and Maunsell, 1999; McAdams and Maunsell, 2000; Moran and Desimone, 1985; Motter, 1993; Reynolds et al., 2000). When using small stimulus sets, tuning shifts that do not align closely with the dimension being probed might be missed. In fact, a shift along orthogonal, unprobed dimensions will appear as a change in response baseline or gain. Our study was designed to increase the likelihood of identifying tuning shifts by measuring responses to a large number of multidimensional stimuli and by characterizing tuning to both orientation and spatial frequency.

Despite the greater sensitivity to tuning shifts provided by our technique, the magnitude and frequency of spectral tuning shifts reported here represent a lower bound on the true magnitude and frequency with which they occur during natural vision. The SRF provides an effective, second-order model of neuronal shape selectivity in area V4 (David et al., 2006), but V4 neurons are also tuned to features not captured by the SRF (Cadieu et al., 2007; Desimone and Schein, 1987; Gallant et al., 1996). Tuning along these unmodeled dimensions might also be modulated by attention, but they would appear to be changes in response baseline or gain. Thus, our measurements of baseline and gain modulation may partially reflect changes in tuning outside the scope of the SRF. Our study also examined only a relatively small number of feature-based attention conditions, compared to the vast number of possible conditions, and target features were chosen independently of the tuning preferences of the neurons being studied. Exploration of a wider range of attention states and tailoring target features to be at or near preferred tuning is likely to reveal additional tuning modulation.

standard error of the mean in each bin. Responses are enhanced during T1 trials only for stimuli that elicit strong responses under both feature-based attention conditions. Green curve gives the difference in response to distractors when spatial attention is directed into and away from the receptive field, as a function of the average response ($n = 87$). Responses are enhanced when attention is directed into the receptive field, regardless of the effectiveness of the stimuli.

Tuning Shifts May Be Conferred by Gain Changes in More Peripheral Visual Areas

Most previous neurophysiological studies of attention in V4 have reported modulation of baseline, gain, and contrast response (Desimone and Duncan, 1995; McAdams and Maunsell, 1999; Reynolds et al., 2000; Williford and Maunsell, 2006). These findings are consistent with the idea that neurons in sensory cortex function as labeled lines that encode input features consistently, regardless of the state of attention. Such modulation is well described by the *feature-similarity gain model* (Maunsell and Treue, 2006; Treue and Martinez-Trujillo, 1999), which holds that neurons increase their gain when attention is directed to their preferred feature or location. However, the feature-similarity gain model does not account for the changes we observe in the spectral tuning profiles of individual neurons that violate the constraints of a labeled line.

Modulation of spectral tuning requires a feedback mechanism capable of adjusting the effective weights of synapses that provide visual input to V4 neurons. One computational model proposes that tuning shifts are mediated by top-down feedback signals that increase the gain of neurons whose tuning matches the target of attention (Compte and Wang, 2006). (This pattern of gain change is similar to that proposed by the feature-similarity gain model.) Increasing the gain of this specific subpopulation of neurons will cause downstream neurons to shift their tuning toward the target of attention. The gain changes we observe for both feature-based and spatial attention are compatible with the first stage of this model. We find that neurons respond more strongly when attention is directed to the preferred feature (i.e., the target eliciting the stronger response) or the preferred spatial location (i.e., into the receptive field, and see [Luck et al., 1997; Martinez-Trujillo and Treue, 2004; McAdams and Maunsell, 1999; Mehta et al., 2000]). The shifts we observe in spectral tuning are compatible with the second stage of this model. Neurons tend to shift their tuning to match the attended feature, as predicted for the downstream neural population if response gain is enhanced in the subset of input neurons that prefer the attended feature.

The tuning shifts effected by feature-based attention occur independently of spatial attention (Hayden and Gallant, 2005; McAdams and Maunsell, 2000; Rossi and Paradiso, 1995; Saenz et al., 2003; Treue and Martinez-Trujillo, 1999), suggesting that these two forms of attention operate through different feedback networks (Maunsell and Treue, 2006). Area V4 is organized retinotopically, and many previous studies have shown that spatial attention only modulates responses near a single retinotopic location (Desimone and Duncan, 1995; McAdams and Maunsell, 1999; Reynolds et al., 2000; Williford and Maunsell, 2006). The distribution of neurons tuned for different spectral features within area V4 is not known, but it is likely that each spectral feature is represented by an anatomically distributed set of neurons. Feedback signals for feature-based attention that modulate tuning must somehow target just those neurons that represent relevant spectral features.

Area V4 and the Matched Filter Hypothesis

Our data suggest that area V4 can function as an attention-mediated matched filter that discards irrelevant information about

the stimulus and enhances the representation of information most relevant to the task at hand. This idea is consistent with proposals from computational modeling studies (Compte and Wang, 2006; Tsotsos et al., 1995) and with Kalman filtering schemes for signal detection (Rao and Ballard, 1997). In the most extreme theoretical case, individual neurons could act as matched filters and shift their tuning to match the target exactly. However, our data show that feature-based attention does not completely change the tuning properties of neurons in V4, but merely shifts tuning *toward* the attended feature. This is not unexpected given the finite number of synaptic connections providing input to each V4 neuron, and the relatively fast timescale of attentional modulation (Olshausen et al., 1993). Therefore, single V4 neurons are not perfect matched filters, but a population of these neurons could function together as a matched filter (or could contribute to a complete matched filter at more central stages of processing).

The effects of spatial attention in V4 are also compatible with a matched filter. If spatial attention engages a matched filter, then the spatial envelope of receptive fields should shift toward the attended location without changing spectral tuning. Previous studies have demonstrated that spatial attention can indeed modulate the spatial receptive field envelope (Connor et al., 1996, 1997; Tolias et al., 2001), and we report here that spatial attention has little effect on spectral tuning. Taken together, these results suggest that shifts might occur only along feature and spatial dimensions only when they match, respectively, the target of feature-based and spatial attention.

Although we did observe shifts in spectral tuning in the FVVS data, we did not observe significant evidence for a matched filter. This difference from the MTS data likely reflects the fact that the signal-to-noise level of the MTS data is much higher than that of FVVS (because of the difficulty of tracking eye position with complete accuracy). However, we cannot rule out the possibilities that this difference instead reflects a strategy unique to free-viewing visual search or the influence of eye movement control signals that attenuate the effects of feature-based attention.

The experiments reported here used targets and distractors selected from a pool of complex natural images. In this paradigm, the target can be identified by one of many features that distinguish it from the distractors, and a matched filter could shift tuning toward any of these distinct features. During MTS, matched filter effects were stronger along the dimension of spatial frequency, which suggests that subjects were attending preferentially to the spatial frequency spectrum of the target. The tendency to shift spatial frequency tuning rather than orientation tuning may simply reflect the strategy used during this specific task, or it might reflect a more fundamental constraint on how spectral tuning can be modulated. This question can be answered with studies using simpler stimuli that more strictly constrain the task-relevant features.

Conclusion

The functional properties of V4 neurons are typically described in terms of static, feed-forward models compatible with the labeled-line hypothesis (Cadiou et al., 2007; David et al., 2006; Pasupathy and Connor, 1999). Our findings suggest that such

models are incomplete and that the tuning properties of V4 neurons change dynamically to meet behavioral demands. Area V4 is critically involved in intermediate stages of visual processing, such as figure-ground segmentation, grouping, and pattern recognition (Gallant et al., 2000; Schiller and Lee, 1991), and it is likely that dynamic tuning shifts in V4 play a critical role in these processes. Just as individual V4 neurons participate in representing visual objects by decomposing them into different dimensions, they participate in attentional selection by modulating their tuning along those dimensions.

EXPERIMENTAL PROCEDURES

Data Collection

All procedures were in accordance with the NIH Guide for the Care and Use of Laboratory Animals and approved by the Animal Care and Use Committee at the University of California, Berkeley. Single-neuron activity was recorded from area V4 of four macaques (*Macaca mulatta*), two while performing an MTS task (Hayden and Gallant, 2005) and the other two while performing an FVVS task (Mazer and Gallant, 2003). Behavioral control, stimulus presentation, and data collection were performed using custom software running on Linux microcomputers. Eye movements were recorded using an infrared eye tracker (500 Hz, Eyelink II, SR Research, Osgoode, ON, Canada). Neuronal activity was recorded using high-impedance (10–25 M Ω) epoxy-coated tungsten microelectrodes (Frederick Haer Company, Bowdoin, ME). Single neurons were isolated and spike events were recorded with 1 ms resolution using a dedicated multi-channel recording system (MAP, Plexon, Dallas, TX).

At the beginning of each recording session, spatial receptive fields were determined while each animal performed a fixation task using an automated mapping procedure (Mazer and Gallant, 2003). Reverse correlation was used to generate spatial kernels from responses to sparse noise, which were fit with a two-dimensional Gaussian to estimate RF location (mean) and size (standard deviation).

Match-to-Sample Task

The MTS task required the identification of a specific natural image patch in a stream of natural image patches at one of two spatial locations during fixation (see Figure 1A and Hayden and Gallant, 2005). On each trial, a cue indicated one of two possible target natural images (feature cue) and to a location either within the RF or in the opposite quadrant of the visual field (spatial cue). New target images were chosen each day before beginning neurophysiological recordings, in order to optimize behavioral performance. Therefore, responses to target stimuli varied substantially across the neurons in our sample. Following a delay period, two rapid, randomly ordered sequences of images (~4 Hz) appeared at the two locations. Subjects responded to the reappearance of the target at the cued location by releasing a bar. Frequent spatial and feature catch stimuli were presented to ensure that subjects were performing as expected. For each neuron, we obtained responses to a large set of natural images (600–1200 distinct images) in each of four crossed attention conditions (i.e., two conditions of feature-based attention and two conditions of spatial attention).

Search targets and distractors were circular natural image patches extracted at random from a library of black and white digital photographs (Corel Inc., Eden Prairie, MN). Patches were cropped to the size of the receptive field of each recorded neuron, and the outer 10% of each patch was blended smoothly into the gray background.

Free-Viewing Visual Search Task

The FVVS task required detection of a specific natural image patch embedded in a random array of visually similar distractors, with no constraints on eye movements (see Figure 1B and Mazer and Gallant, 2003). Trials were cued by onset of a textured $1/f^2$ -power random noise pattern that served as the background pattern for the search task. Each trial was initiated when the animal grabbed the capacitive touch bar. A search target was then presented for 2–5 s at the center of the display. Animals were allowed to inspect the target

using voluntary eye movements. After a 2–4 s delay, an array of 9–25 potential match stimuli appeared, and this remained visible for 2–5 s. If the array contained the target, the animal had to release the touch bar within 500 ms after array offset in order to receive a reward. (They were not required to indicate the position of the target, but only its presence.) If the array did not contain the target, then the animal had to continue to hold the bar until another array appeared after a 2–3 s delay. Up to seven different arrays could occur in any single trial. The temporal and spatial position of the target was selected at random before each trial, as were the positions of all distractors.

Search targets and distractors were circular image patches extracted at random from the same library of digital photographs used for MTS. (One photograph was chosen as the image patch source for each recording session.) Patches were cropped to the size of the receptive field of each recorded neuron, and the outer 10% of each patch was blended smoothly into the background pattern. Search array spacing was adjusted so that fixation of one patch placed a different patch close to the center of the receptive field of the recorded cell. For parafoveal neurons (<2° eccentricity), patch size was adjusted so that patches encompassed both the RF and the fovea, and array spacing was adjusted to prevent overlap of the patches.

We did not control spatial attention during FVVS. Instead, we controlled feature-based attention by varying the search target from trial to trial and assumed that spatial attention effects were averaged out over the large number of unconstrained eye movements. SRFs were estimated according to the no modulation, baseline/gain modulation, and tuning modulation models using the same procedure as for the MTS data. Stimuli for SRF estimation were generated by using measured eye movements to reconstruct the spatiotemporal visual input stream that fell in the receptive field during the search phase of the FVVS task.

Characterization of Attention-Dependent Visual Tuning Properties

Neurons in area V4 exhibit several nonlinear response properties, including phase invariance and position invariance (Desimone and Schein, 1987; Gallant et al., 1996). Therefore, to accurately characterize the visual tuning of V4 neurons, we applied a nonlinear spatial Fourier power transformation to each stimulus (David et al., 2006). This transform discards absolute spatial information and makes the stimulus-response relationship more linear (Wu et al., 2006). We call the linear mapping from power-transformed stimulus to neural response the spectral receptive field (SRF). We assessed the effects of attention on neural responses by testing whether the SRF changed with attention conditions.

The *no modulation model* assumes that attention does not have any effect on visual responses. It was fit by estimating the SRF with all the data acquired during behavior, without considering the target identity or spatial position. The relationship between the stimulus and response is described by the SRF, h_0 ,

$$r(t) = \sum_{\omega_x = -N}^N \sum_{\omega_y = -N}^N |\hat{S}(\omega_x, \omega_y, t)|^2 h_0(\omega_x, \omega_y) + d_0$$

where $r(t)$ is the firing rate 50–200 ms after the onset of stimulus frame t , and $\hat{S}(\omega_x, \omega_y, t)$ is the two-dimensional spatial Fourier transform (i.e., Fourier power spectrum) of the stimulus in the receptive field. The constant, d_0 describes the baseline firing rate.

The *baseline/gain modulation model* assumes that attention can change the response baseline and/or gain but that it does not change tuning (i.e., the structure of the SRF). This model was fit by introducing a state variable, a , corresponding to the state of attention (i.e., feature-based attention to target T1 or T2 or spatial attention inside or outside the RF),

$$r(t) = g(a) \sum_{\omega_x = -N}^N \sum_{\omega_y = -N}^N |\hat{S}(\omega_x, \omega_y, t)|^2 h_0(\omega_x, \omega_y) + d(a)$$

where h_0 is the SRF fit from the no modulation model and $d(a)$, baseline response, and $g(a)$, global gain are free to vary with attention state.

The *tuning modulation model* assumes that attention can change the spectral tuning profile of a neuron. In addition to response baseline and gain, the SRF, $h(\omega_x, \omega_y, a)$, is fit separately for each attention state,

$$r(t) = g(a) \sum_{\omega_x = -N}^N \sum_{\omega_y = -N}^N |\hat{S}(\omega_x, \omega_y, t)|^2 h(\omega_x, \omega_y, a) + d(a)$$

Model Fitting Procedure

Each SRF was estimated using a normalized reverse correlation procedure that finds the best linear mapping between stimulus and response (David et al., 2004, 2006; Theunissen et al., 2001). For the tuning modulation model, SRFs were estimated using the same method as for the no modulation model, but the data used to fit each SRF were drawn only from individual attention conditions. SRFs were estimated only from responses to distractors (i.e., excluding target and catch stimuli) recorded during correct trials; responses to targets and responses recorded on error trials were excluded. To eliminate potential bias in SRFs estimated under different attention conditions, the same regularization was applied to all SRFs for a single neuron. In the MTS experiment, where stimuli could be controlled exactly, the same distractors were used in each of the four behavioral conditions, eliminating the possibility of sampling bias.

Our model fitting procedure can be viewed as stepwise regression, because each successively more complicated model encompasses all of the simpler models. For each neuron, we used a crossvalidation procedure to determine whether each model represented a significant improvement over the simpler alternatives. Each model was fit using only 95% of the available data, and the resulting model was used to predict responses in the remaining 5%. Model performance was quantified by computing the squared correlation coefficient (Pearson's r) between predicted and observed responses. Statistical significance of modulation by attention was assessed by a jackknifing procedure in which the estimation-validation analysis was repeated 20 times, each time reserving a different 5% of data for validation. A model was taken to provide significantly improved predictions if its prediction score was significantly greater ($p < 0.05$) than the simpler model.

Significance Tests

Unless otherwise specifically mentioned, we used a one-tailed, jackknifed t test to verify the statistical significance of our findings (Efron and Tibshirani, 1986). In many cases, a traditional t test is sufficient to determine whether two mean values are significantly different. However, the traditional t test assumes that individual measurements follow a Gaussian distribution, and estimates of standard error will be biased if the distributions are not Gaussian. The jackknifed t test uses a bootstrapping procedure that avoids potential bias from non-Gaussian distributions in measurements of standard error.

Target Similarity Index

To determine whether SRFs estimated using the tuning modulation model to shift their tuning toward the target of feature-based attention, we measured a *target similarity index* (TSI) for each neuron. The TSI is the change in correlation (i.e., normalized dot product) between the SRF and the spectra of the two target images on attended versus nonattended trials,

$$TSI = \frac{h_1 \cdot \hat{t}_1 + h_2 \cdot \hat{t}_2 - h_1 \cdot \hat{t}_2 - h_2 \cdot \hat{t}_1}{2 - h_1 \cdot \hat{t}_2 - h_2 \cdot \hat{t}_1}$$

Here, \hat{t}_1 and \hat{t}_2 are the Fourier power spectra of the two target images, T1 and T2, and h_1 and h_2 are SRFs estimated separately under each feature-based attention condition. Both the power spectra and the SRFs are normalized so that their standard deviation is 1. The dot indicates point-wise multiplication and then summing over spatial frequencies. Index values greater than zero indicate SRFs that more closely match the spectrum of T1 on trials when feature-based attention is directed toward T1 and/or more closely match the spectrum of T2 on trials when attention is directed toward T2. A value of 1 indicates a perfect match between SRF and target in both attention conditions.

To study spectral tuning shifts in more detail, we also computed the TSI separately for orientation and spatial frequency tuning. SRFs and target power spectra were decomposed by singular value decomposition into orientation and spatial frequency curves that best predicted the full two-dimensional spectrum (David et al., 2006). The same equation was then used to compute TSI values separately for each of these tuning curves.

Data Preprocessing

To prepare visual stimuli for analysis, the stimulus falling in the receptive field was downsampled to 20×20 pixels, multiplied by a Hanning window with radius equal to the receptive field (to reduce edge artifacts), and transformed into the Fourier power domain by computing and squaring the two-dimensional FFT (David et al., 2006). The response to each stimulus frame was taken as the spike rate (spikes/s) 50–200 ms after stimulus onset.

To analyze data acquired in the FVVS task, we first had to determine what visual stimuli fell in and around the receptive field during each fixation. Eye calibration data (Mazer and Gallant, 2003) were used to determine where each animal was looking at every moment. Fixations and saccades were distinguished by smoothing and differentiating measured eye velocity; a fixation was defined as any period where the eye remained stationary for 100 ms. Eye movement statistics for an example behavior session appear in Figure S2.

SUPPLEMENTAL DATA

Supplemental Data include three figures and can be found with this article online at <http://www.neuron.org/cgi/content/full/59/3/509/DC1>.

ACKNOWLEDGMENTS

This work was supported by grants from the National Science Foundation (S.V.D.) and the National Eye Institute (J.L.G.). The authors would like to thank Kate Gustavsen, Kathleen Hansen, and two anonymous reviewers for helpful comments on the manuscript.

Accepted: July 2, 2008

Published: August 13, 2008

REFERENCES

- Adrian, E.D., and Matthews, R. (1927). The action of light on the eye: Part I. The discharge of impulses in the optic nerve and its relation to the electric changes in the retina. *J. Physiol.* **63**, 378–414.
- Barlow, H.B. (1972). Single units and sensation: a neuron doctrine for perceptual psychology? *Perception* **1**, 371–394.
- Cadiou, C., Kouh, M., Pasupathy, A., Connor, C., Riesenhuber, M., and Poggio, T.A. (2007). A model of V4 shape selectivity and invariance. *J. Neurophysiol.* **98**, 1733–1750.
- Carrasco, M., Ling, S., and Read, S. (2004). Attention alters appearance. *Nat. Neurosci.* **7**, 308–313.
- Compte, A., and Wang, X.J. (2006). Tuning curve shift by attention modulation in cortical neurons: a computational study of its mechanisms. *Cereb. Cortex* **16**, 761–778.
- Connor, C.E., Gallant, J.L., Preddie, D.C., and Van Essen, D.C. (1996). Responses in area V4 depend on the spatial relationship between stimulus and attention. *J. Neurophysiol.* **75**, 1306–1308.
- Connor, C.E., Preddie, D.C., Gallant, J.L., and Van Essen, D.C. (1997). Spatial attention effects in macaque area V4. *J. Neurosci.* **17**, 3201–3214.
- David, S.V., and Gallant, J.L. (2005). Predicting neuronal responses during natural vision. *Network* **16**, 239–260.
- David, S.V., Vinje, W.E., and Gallant, J.L. (2004). Natural stimulus statistics alter the receptive field structure of V1 neurons. *J. Neurosci.* **24**, 6991–7006.
- David, S.V., Hayden, B.Y., and Gallant, J.L. (2006). Spectral receptive field properties explain shape selectivity in area V4. *J. Neurophysiol.* **96**, 3492–3505.
- Desimone, R., and Schein, S.J. (1987). Visual properties of neurons in area V4 of the macaque: Sensitivity to stimulus form. *J. Neurophysiol.* **57**, 835–868.
- Desimone, R., and Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* **18**, 193–222.

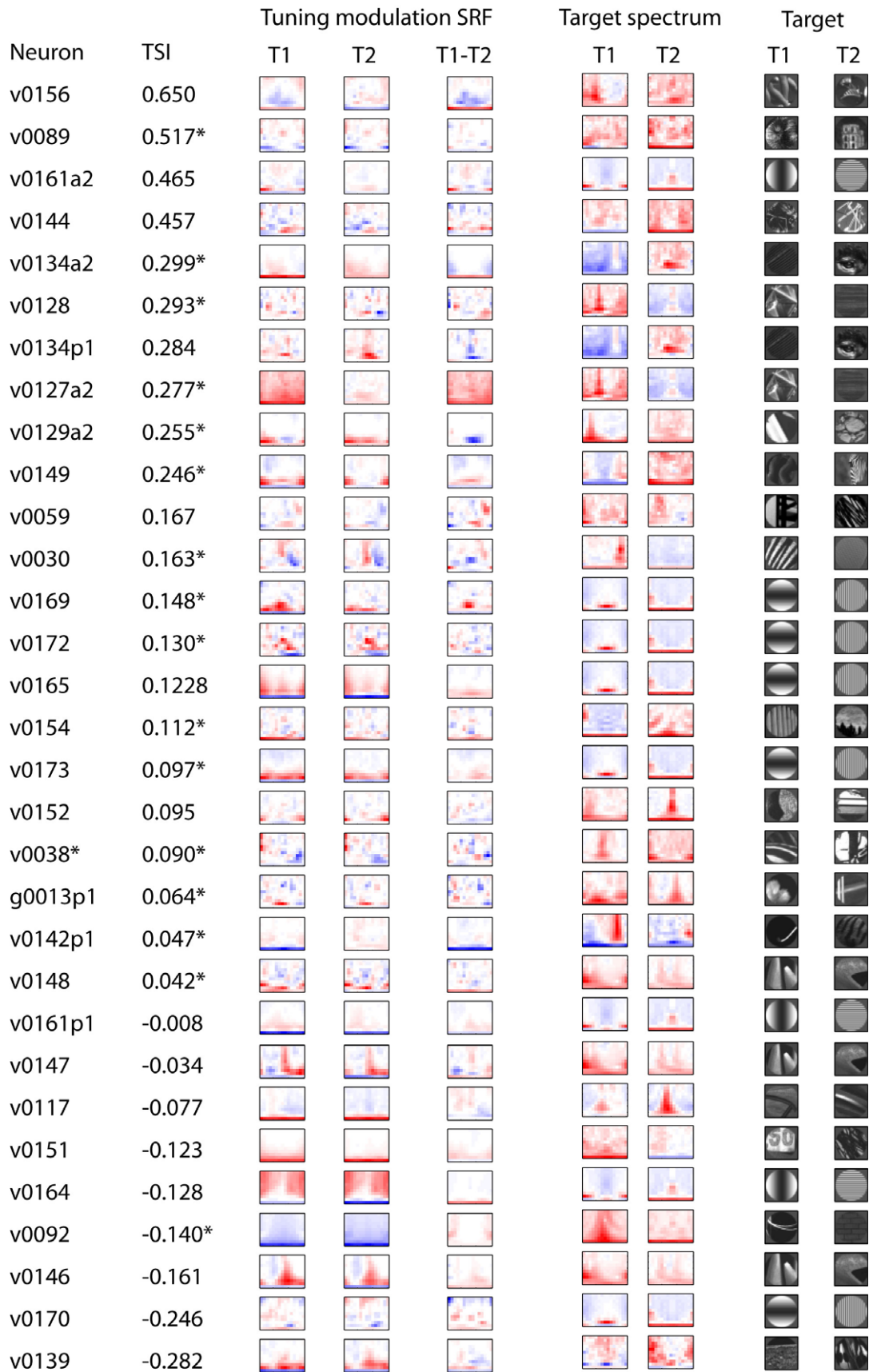
- Efron, B., and Tibshirani, R. (1986). Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Stat. Sci.* **1**, 54–75.
- Fallah, M., Stoner, G.R., and Reynolds, J.H. (2007). Stimulus-specific competitive selection in macaque extrastriate visual area V4. *Proc. Natl. Acad. Sci. USA* **104**, 4165–4169.
- Gallant, J.L., Braun, J., and Van Essen, D.C. (1993). Selectivity for polar, hyperbolic, and Cartesian gratings in macaque visual cortex. *Science* **259**, 100–103.
- Gallant, J.L., Connor, C.E., Rakshit, S., Lewis, J.W., and Van Essen, D.C. (1996). Neural responses to polar, hyperbolic, and cartesian gratings in area V4 of the macaque monkey. *J. Neurophysiol.* **76**, 2718–2739.
- Gallant, J.L., Shoup, R.E., and Mazer, J.A. (2000). A human extrastriate cortical area functionally homologous to Macaque V4. *Neuron* **27**, 227–235.
- Gattass, R., Souza, A.P.B., and Gross, C.G. (1988). Visuotopic organization and extent of V3 and V4 of the macaque. *J. Neurosci.* **8**, 1831–1845.
- Gilbert, C.D., and Sigman, M. (2007). Brain states: top-down influences in sensory processing. *Neuron* **54**, 677–696.
- Haenny, P.E., and Schiller, P.H. (1988). State dependent activity in monkey visual cortex. 1. Single cell activity in V1 and V4 on visual tasks. *Exp. Brain Res.* **69**, 225–244.
- Haenny, P.E., Maunsell, J.H.R., and Schiller, P.H. (1988). State dependent activity in monkey visual cortex. II. Retinal and extraretinal factors in V4. *Exp. Brain Res.* **69**, 245–259.
- Hayden, B.Y., and Gallant, J.L. (2005). Time course of attention reveals different mechanisms for spatial and feature-based attention in area V4. *Neuron* **47**, 637–643.
- Lee, D.K., Itti, L., Koch, C., and Braun, J. (1999). Attention activates winner-take-all competition among visual filters. *Nat. Neurosci.* **2**, 375–381.
- Lu, Z.L., and Doshier, B.A. (2004). Spatial attention excludes external noise without changing the spatial frequency tuning of the perceptual template. *J. Vis.* **4**, 955–966.
- Luck, S.J., Chelazzi, L., Hillyard, S.A., and Desimone, R. (1997). Neural mechanisms of spatial selective attention in areas V1, V2 and V4 of macaque visual cortex. *J. Neurophysiol.* **77**, 24–42.
- Marr, D. (1982). *Vision* (San Francisco: W.H. Freeman).
- Martinez-Trujillo, J.C., and Treue, S. (2004). Feature-based attention increases the selectivity of population responses in primate visual cortex. *Curr. Biol.* **14**, 744–751.
- Maunsell, J.H., and Cook, E.P. (2002). The role of attention in visual processing. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **357**, 1063–1072.
- Maunsell, J.H., and Treue, S. (2006). Feature-based attention in visual cortex. *Trends Neurosci.* **29**, 317–322.
- Mazer, J.A., and Gallant, J.L. (2003). Goal-related activity in V4 during free viewing visual search. Evidence for a ventral stream visual salience map. *Neuron* **40**, 1241–1250.
- Mazer, J.A., Vinje, W.E., McDermott, J., Schiller, P.H., and Gallant, J.L. (2002). Spatial frequency and orientation tuning dynamics in area V1. *Proc. Natl. Acad. Sci. USA* **99**, 1645–1650.
- McAdams, C.J., and Maunsell, J.H.R. (1999). Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *J. Neurosci.* **19**, 431–441.
- McAdams, C.J., and Maunsell, J.H.R. (2000). Attention to both space and feature modulates neuronal responses in macaque area V4. *J. Neurophysiol.* **83**, 1751–1755.
- Mehta, A.D., Ulbert, I., and Schroeder, C.E. (2000). Intermodal selective attention in monkeys. I: distribution and timing of effects across visual areas. *Cereb. Cortex* **10**, 343–358.
- Moran, J., and Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science* **229**, 782–784.
- Motter, B.C. (1993). Focal attention produces spatially selective processing in visual cortical areas V1, V2 and V4 in the presence of competing stimuli. *J. Neurophysiol.* **70**, 909–919.
- Motter, B.C. (1994). Neural correlates of attentive selection for color or luminance in extrastriate area V4. *J. Neurosci.* **14**, 2178–2189.
- Ogawa, T., and Komatsu, H. (2004). Target selection in area V4 during a multi-dimensional visual search task. *J. Neurosci.* **24**, 6371–6382.
- Olshausen, B.A., Anderson, C.H., and Van Essen, D.C. (1993). A neurobiological model of visual attention and invariant pattern recognition based on dynamic rerouting of information. *J. Neurosci.* **13**, 4700–4719.
- Pasupathy, A., and Connor, C.E. (1999). Responses to contour features in macaque area V4. *J. Neurophysiol.* **82**, 2490–2502.
- Rao, R.P.N., and Ballard, D.H. (1997). Dynamic model of visual recognition predicts neural response properties in the visual cortex. *Neural Comput.* **9**, 721–763.
- Reynolds, J.H., and Chelazzi, L. (2004). Attentional modulation of visual processing. *Annu. Rev. Neurosci.* **27**, 611–647.
- Reynolds, J.H., Pasternak, T., and Desimone, R. (2000). Attention increases sensitivity of V4 neurons. *Neuron* **26**, 703–714.
- Rossi, A.F., and Paradiso, M.A. (1995). Feature-specific effects of selective visual attention. *Vision Res.* **35**, 621–634.
- Saenz, M., Buracas, G.T., and Boynton, G.M. (2003). Global feature-based attention for motion and color. *Vision Res.* **43**, 629–637.
- Schiller, P.H., and Lee, K. (1991). The role of primate extrastriate area V4 in vision. *Science* **251**, 1251–1253.
- Theunissen, F.E., David, S.V., Singh, N.C., Hsu, A., Vinje, W.E., and Gallant, J.L. (2001). Estimating spectro-temporal receptive fields of auditory and visual neurons from their responses to natural stimuli. *Network* **12**, 289–316.
- Tolias, A.S., Moor, T., Smirnakis, S.M., Tehovnik, E.J., Siapas, A.G., and Schiller, P.H. (2001). Eye movements modulate visual receptive fields of V4 neurons. *Neuron* **29**, 757–767.
- Treue, S., and Martinez-Trujillo, J.C. (1999). Feature-based attention influences motion processing gain in macaque visual cortex. *Nature* **399**, 575–579.
- Tsotsos, J.M., Culhane, S.M., Wai, W.Y.K., Lai, Y., Davis, N., and Nufflo, F. (1995). Modeling visual attention via selective tuning. *Artif. Intell.* **78**, 507–545.
- Williford, T., and Maunsell, J.H. (2006). Effects of spatial attention on contrast response functions in macaque area V4. *J. Neurophysiol.* **96**, 40–54.
- Womelsdorf, T., Anton-Erxleben, K., Pieper, F., and Treue, S. (2006). Dynamic shifts of visual receptive fields in cortical area MT by spatial attention. *Nat. Neurosci.* **9**, 1156–1160.
- Woolley, S.M., Fremouw, T.E., Hsu, A., and Theunissen, F.E. (2005). Tuning for spectro-temporal modulations as a mechanism for auditory discrimination of natural sounds. *Nat. Neurosci.* **8**, 1371–1379.
- Wu, M.C.-K., David, S.V., and Gallant, J.L. (2006). Complete functional characterization of sensory neurons by system identification. *Annu. Rev. Neurosci.* **12**, 477–505.

Neuron, Volume 59

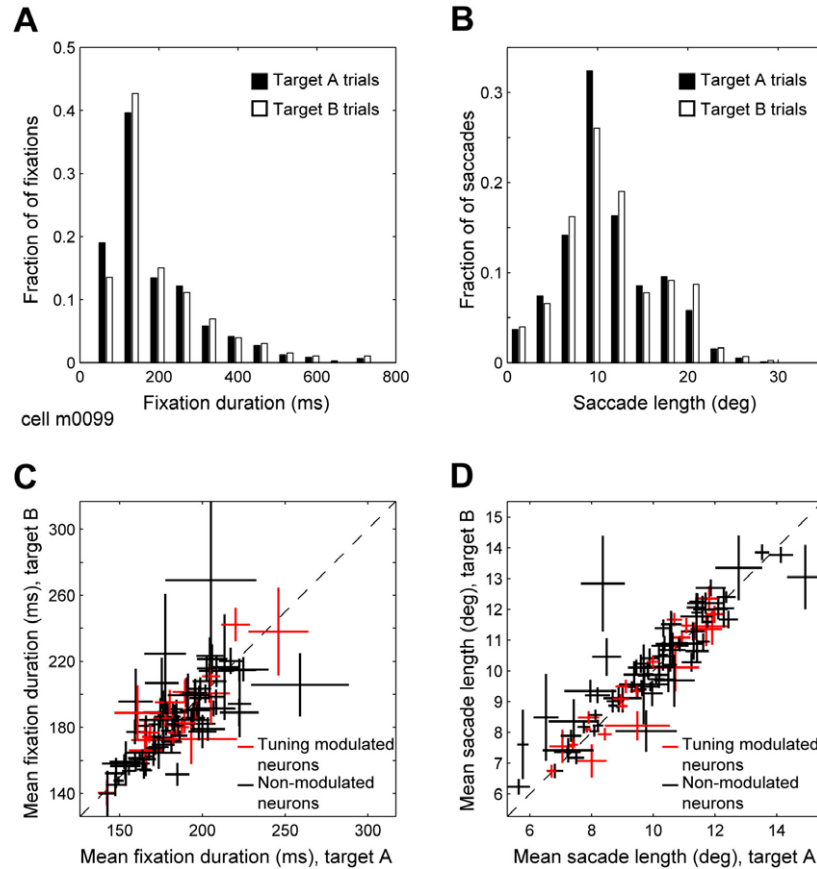
Supplemental Data

**Attention to Stimulus Features Shifts Spectral
Tuning of V4 Neurons during Natural Vision**

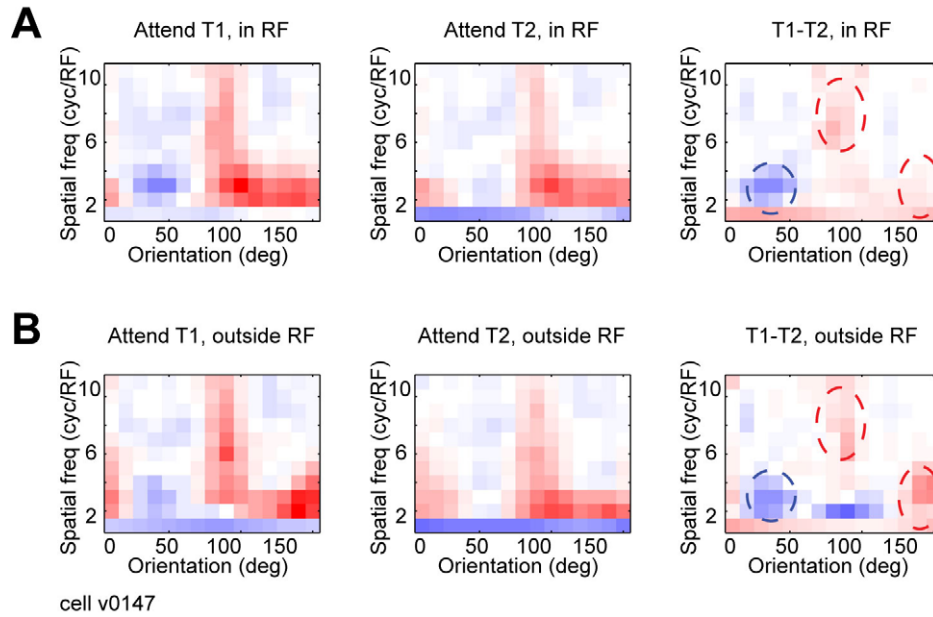
Stephen V. David, Benjamin Y. Hayden, James A. Mazer, and Jack L. Gallant



Supplementary Figure 1. Details of matched filter effects during MTS. The three columns labeled *Tuning modulation SRF* show SRFs for all neurons (31/105) that show significant tuning modulation estimated during different conditions of feature-based attention (i.e., when the target was either T1 or T2), and their difference. The two columns labeled *Target spectrum* show the Fourier power spectra of targets T1 and T2 used for each neuron. Here the average power spectrum of all distractors has been subtracted to highlight spectral features that are either stronger (red) or weaker (blue) than average. The two columns labeled *Target* show the original T1 and T2 images. The column labeled *TSI* gives the target similarity index for each neuron (for details see Methods). TSI values significantly different from zero are marked with an asterisk (* $p < 0.05$, jackknifed t -test).



Supplementary Figure 2. Eye movement statistics during FVVS are not correlated with spectral tuning shifts. **A**, Histogram comparing the duration of fixations during target A trials (black) and target B trials (white) during recording from the neuron in Figure 7. The average fixation duration (A trials, 190 ms; B trials, 201 ms) is not significantly different between target conditions (jackknifed t -test). **B**, Histogram comparing the length of saccades during target A trials (black) and target B trials (white) during recording from the same neuron. The average saccade length (A trials, 11 deg; B trials, 11.5 deg) is not significantly different between target conditions (jackknifed t -test). **C**, Comparison of average fixation duration between target conditions for all FVVS neurons. Each cross indicates the mean and standard error during target A (horizontal) and target B (vertical) conditions. The average difference between conditions for neurons with significant tuning modulation (red, 10.3 ms) is not significantly different than for neurons that are not modulated (black, 10.9 ms, jackknifed t -test). **D**, Comparison of average saccade length between target conditions for all FVVS neurons. Data is presented in the same format as C. The average difference in saccade length between conditions for neurons with significant tuning modulation (0.49 deg) is not significantly different than for neurons that are not modulated (0.54 deg, jackknifed t -test).



Supplementary Figure 3. Tuning modulation by feature-based attention occurs independently of spatial attention. **A**, SRFs estimated for a single neuron when feature-based attention was directed either toward target T1 (left) or T2 (middle), using only data from trials when spatial attention was directed into the neuron’s receptive field (see Supplementary Figure 1 for details about targets). The difference between SRFs (right) shows complex tuning modulation, increasing gain at some orientations and spatial frequencies (red ovals) while decreasing gain at others (blue oval). **B**, SRFs estimated for the same neuron under each feature-based attention condition, but only using trials when spatial attention was directed away from the receptive field. The difference between SRFs reveals a similar tuning shift.