Ameliorating the combinatorial explosion with spatial frequency-matched combinations of V1 outputs

Sarah Hancock,
Nottingham Visual Neuroscience, School of Psychology, University of Nottingham, Nottingham, UK

David P. McGovern, and
Nottingham Visual Neuroscience, School of Psychology, University of Nottingham, Nottingham, UK

Jonathan W. Peirce
Nottingham Visual Neuroscience, School of Psychology, University of Nottingham, Nottingham, UK

Abstract

Little is known about the way in which the outputs of early orientation-selective neurons are combined. One particular problem is that the number of possible combinations of these outputs greatly outweighs the number of processing units available to represent them. Here we consider two of the possible ways in which the visual system might reduce the impact of this problem. First, the visual system might ameliorate the problem by collapsing across some low-level feature coded by previous processing stages, such as spatial frequency. Second, the visual system may combine only a subset of available outputs, such as those with similar receptive field characteristics. Using plaid-selective contrast adaptation and the curvature aftereffect, we found no evidence for the former solution; both aftereffects were clearly tuned to the spatial frequency of the adaptor relative to the test probe. We did, however, find evidence for the latter with both aftereffects; when the components forming our compound stimuli were dissimilar in spatial frequency, the effects of adapting to them were substantially reduced. This has important implications for mid-level visual processing, both for the combinatorial explosion and for the selective "binding" of common features that are perceived as coming from a single visual object.

Keywords

middle vision; perceptual organization; shape and contour; spatial vision

Introduction

The basic function of very many neurons throughout mammalian cortex is to combine the outputs of other neurons and provide a transformed representation of that combined signal. Often in the visual system this is performed in a feedforward hierarchy whereby neurons

combine the outputs of some previous level in the hierarchy, frequently also taking account of additional feedback information from higher order areas (Felleman & Van Essen, 1991). For example, the outputs of retinal photoreceptors are combined by retinal ganglion cells to form center–surround receptive fields, capable of detecting local contrast. In a similar manner, the outputs of cells in the Lateral Geniculate Nucleus (LGN) are combined by neurons in the Primary Visual Cortex (V1) to form elongated receptive fields capable of detecting orientation. This combination might be a linear one, in which the final signal is directly proportional to the summed outputs of the input signals, or highly nonlinear (e.g., Peirce, 2007a).

Rather little is known about how the outputs of V1 neurons are combined and what form of receptive field might follow. Potentially, responses from V1 with different preferred orientations could be summed to form contours of various curvatures. Certainly it has been shown that V2 neurons encode combinations of different orientations within their receptive field (Anzai, Peng, & Van Essen, 2007) and that neurons in V4 appear to be strongly sensitive to the presence of particular curved contours on their receptive fields (Gallant, Braun, & Van Essen, 1993; Gallant, Connor, Rakshit, Lewis, & Van Essen, 1996; Pasupathy & Connor, 1999, 2001, 2002). Additionally, psychophysical studies have shown an aftereffect generated by curved contours (Gheorghiu & Kingdom, 2006, 2007) and that this is greater than predicted by the local tilt aftereffects of the components (Hancock & Peirce, 2008).

For V1 cells with overlapping receptive fields of different orientations, the combined outputs might form mechanisms responding selectively to plaids. Although no electrophysiological studies have so far reported neurons for which the preferred stimulus was a plaid, it is clear that neurons in MT/V5 are capable of responding to the overall perceived motion of a plaid, rather than its parts (e.g., Movshon, Adelson, Gizzi, & Newsome, 1985; Rust, Mante, Simoncelli, & Movshon, 2006). Again, there is also psychophysical evidence to suggest that adaptation to a plaid is greater than the sum of adaptations to the parts, consistent with such a mechanism (McGovern & Peirce, 2010; Peirce & Taylor, 2006).

There is a problem, however, with the notion that at some stage beyond V1, its outputs are combined into such curvature and plaid detectors, which is that there are far too many pairwise combinations of these cells' outputs for them all to be represented; the well-known problem of "combinatorial explosion." If the set of possible V1 neurons were governed by N parameters (e.g., preferred spatial frequency, orientation bandwidth, etc.) and contained M cells, then the set of pairwise combinations has $2 \times N$ parameters and requires M^2 cells. The problem is accentuated by the fact that V1 is the largest visual area in the brain (Felleman & Van Essen, 1991). The question is how the visual system handles this problem. Here, we examine two of the potential methods that might ameliorate its effects. The first is that compound detectors might be invariant to some of the parameters of the previous processing stages. For instance a curvature detector might respond to a particular curve but across all possible spatial frequencies, such that spatial frequency, as a parameter of the curve, is removed. This would allow many combinations of signals to be represented by a single conjunction detector. The second is that the visual system could build compound detectors from only a subset of the V1 outputs. For example, combining only components that *match each other* in spatial frequency dramatically reduces the number of combinations possible. Furthermore this might carry the additional benefit of limiting detection to edges that appear to come from the same object. These solutions to the problem are independent. It would be possible, for instance, to have a detector that responds well to any spatial frequency of contour provided that spatial frequency was constant along the contour. The converse would also be possible; a detector that requires specific spatial frequencies along its contour but is

constructed to allow these to differ from each other. The same is true for conjunctions that form plaids.

It is clear that in some circumstances, where the combination of different spatial frequencies is informative, mechanisms do exist that seem capable of combining them. For instance, the detection of collinear, overlapping Fourier energy at multiple spatial frequencies can be used to identify features such as hard edges or lines, depending on the relative spatial phase of those energies (Marr & Hildreth, 1980; Morrone & Burr, 1988; Watt & Morgan, 1985). Numerous studies have shown that there are nonlinear interactions between such collinear spatial frequency channels (e.g., Henning, Hertz, & Broadbent, 1975; Morgan & Watt, 1997). Several studies from Olzak and Thomas (1991, 1992) point to the existence of two distinct mechanisms for combining these low-level signals; one that combines across spatial frequencies within a very limited range of orientations, which might mediate the detection of oriented edges, and another that combines across orientations within a limited range of spatial frequencies that they suggest might provide information about texture (Olzak & Thomas, 1999). Similarly, Georgeson and Meese have shown that, although both spatial frequency combinations and orientation combinations can be represented, they seem not to be represented simultaneously (Georgeson, 1998; Georgeson & Meese, 1997).

Here we consider mechanisms that combine energies at different orientations and/or spatial locations, namely putative detectors for plaids and curves. For these there is no clear utility in combining information over different spatial frequencies. The question is whether this nonetheless occurs. We examine whether the effects of plaid-selective and curvature-selective adaptation are invariant to spatial frequency, by adapting participants to stimuli at a fixed spatial frequency and testing with a range of probes. To quantify the effect of selective combination, we measure the adaptation effects with stimuli that are comprised of two components that differ in spatial frequency, but where the adaptor and probe are the same.

General methods

All the experiments in this paper used a "compound adaptation" paradigm, designed to measure adaptation to a compound stimulus beyond that predicted by adaptation to its constituent parts. Two patches on opposite sides of the visual field are adapted simultaneously—one to a compound stimulus consisting of two gratings presented together (compound field), and one to the same two grating stimuli presented in isolation, alternating every second (component field). A test stimulus is then presented in both adapted locations and the point of subjective equality is determined. As the same component gratings are presented in both adapting locations and the total presentation time and the alternation rates for each grating are matched, any aftereffect due to adaptation to the components alone should be equal on both sides. Therefore, any residual difference in the adaptation effect between the two sides must be due to adaptation to the compound as a whole.

Two different forms of compound adaptation were examined. In the first (used in Experiments 1a and 2a), the compound stimulus was comprised from two fully overlapping gratings each at half the maximum contrast, giving rise to a full contrast plaid. Adaptation to this stimulus results in a decrease in the apparent contrast of a subsequently presented test probe. In the second (Experiments 1b and 2b), the compound stimulus was comprised of two gratings that are presented adjacent to each other to form a chevron-like contour. Adaptation to this stimulus leads to a straight test contour appearing curved in the opposite direction. Experiment 1 examined the effect of varying the spatial frequency of the test stimulus relative to the adaptor in order to obtain spatial frequency tuning functions of the underlying mechanisms. In Experiment 2, the test stimulus always matched the adaptor, but the constituent gratings varied relative to each other. This manipulation, in moving plaids, is

known to result in the percept of a pair of translucent gratings rather than that of a coherent plaid pattern (e.g., Adelson & Movshon, 1982) and presumably also reduces the percept of a fully coherent contour.

Participants

Participants consisted of eight healthy volunteers (two experienced observers and six naive participants) with normal or corrected-to-normal vision, who gave their consent. Of these, experienced observers SH and DM took part in all experiments; CS, SXI, and MK participated in the plaid adaptation experiments, and LS, PB, and SQ participated in the contour experiments. All procedures were approved by the School of Psychology Ethics Committee, University of Nottingham, UK.

Apparatus

For the plaid adaptation experiments (Experiments 1a and 2a), stimuli were presented on a computer-controlled cathode-ray-tube (CRT) monitor (Vision Master Pro 454, liyama) at a resolution of 1152×864 pixels and at a refresh rate of 85 Hz with a mean luminance of 108.3 cd/m^2 The observer's head was stabilized in a chin rest 57 cm from the monitor with the viewable area subtending 40.5° visual angle. For the contour adaptation experiments (Experiments 1b and 2b), stimuli were also presented on a CRT monitor (Vision Master Pro 513, liyama) but at a resolution of 1024×768 pixels with a mean luminance of 49.56 cd/m^2. The viewing distance was 52 cm from the monitor giving a viewable area of 43.6° of visual angle.

Both monitors were driven by 14-bit digital-to-analog converters (Bits++, Cambridge Research Systems, Cambridge, UK). Stimuli were presented and data collected using the PsychoPy stimulus generation library (Peirce, 2007b). They were calibrated using a photo-spectrometer (PR650, Photo Research, Chatsworth, CA, USA) to gamma-correct the red, green, and blue (RGB) guns independently and the gamma correction was verified psychophysically using a second-order motion-nulling procedure (Ledgeway & Smith, 1994).

Plaid adaptation

Stimuli—Plaids were constructed from the linear combination of two luminance-modulated sinusoidal gratings at oblique angles, ± 45° from vertical. Each grating contributed equal contrast to the plaid (50% of the maximum contrast of 0.96 Michelson). All stimuli were presented in a Gaussian envelope with a standard deviation of 0.5° visual angle (such that the stimulus had a diameter of 8° at the point where it fell below 1% contrast). The spatial phase of the stimulus was randomly jittered (every 200 ms) across time to prevent retinal afterimages.

Procedure—The basic procedure is shown schematically in Figure 1. Participants were adapted to a pair of component gratings at different locations on the retina, centered at 6° visual angle either side of the fovea on the horizontal meridian. During adaptation both gratings were presented simultaneously as a full contrast plaid in one visual hemi-field (compound field) and individually as two alternating half-contrast gratings, in the other hemi-field (component field). Both total exposure time and temporal frequency were equated for each component grating in the two hemi-fields by having the components alternating every second and the plaid alternating with a blank field every second. The temporal phases of the alternations in each hemi-field were independently randomized. After adaptation, participants compared the contrast of a plaid probe at the same retinal location that it had itself been adapted (the test probe) with one in the opposite location (the reference probe) and were required to report which side had the higher apparent contrast. The

reference probe took a fixed contrast value of 0.42, while the contrast of the test probe gradually decreased or increased in steps using an adaptive 1-up, 1-down staircase procedure designed to maintain stimulus presentation near the point of subjective equality (PSE). Each staircase consisted of 50 test presentations of the probe stimuli.

The initial period of adaptation lasted for 30 s and was "topped-up" with another 2 s of adaptation prior to each trial. This was followed by a 200 ms inter-stimulus interval (ISI) before presentation of the probe stimuli for 200 ms. A fixation spot was visible for the entire trial. Observers pressed one of two keys to make a 2AFC response indicating the side on which the stimulus appeared to have higher contrast, triggering the next trial to commence with another "top-up" adaptation period.

The adaptation procedure was counterbalanced to avoid any side bias, so observers were adapted in separate sessions to trials with the compound adaptor on the right side of fixation and trials where the compound adaptor was on the left. To prevent crossover adaptation between conditions a minimum time of 1 h (and typically much longer) was left between sessions. Each observer collected a minimum of four staircases for each side and stimulus condition.

Contour adaptation

Stimuli—Contour stimuli were constructed from two luminance-modulated sinusoidal gratings oriented to form an oblique, V-shaped contour of 140°. In their original paper, Hancock and Peirce (2008) used as components two partially overlapping Gabor patches to create a continuous contour. Using this method, the range of orientations that can be presented without creating an artifact in the center of the stimulus is limited. To avoid this problem, the stimuli in this study abutted the component gratings along a hard edge and then the entire stimulus was masked by a Gaussian profile. This Gaussian envelope had a standard deviation of 0.5° visual angle, such that the entire stimulus had a diameter of 5° at the point where it fell below 1% contrast. The result is that each component is effectively semi-circular but with a Gaussian profile to the curved edge. The spatial phase of the two gratings was aligned resulting in continuous contours with a more flexible degree of curvature. The stimuli were always presented at the maximum contrast of 0.98 Michelson for this monitor.

The reference probe was a similar oblique contour with a fixed angle of 160° while the test probe varied in contour angle. The center of the contour (or the hard edge of each component) was located 6.5° of visual angle from fixation.

Procedure—The procedure for contour adaptation was very similar to that for the plaid adaptation described above. The initial adaptation period was 60 s. Otherwise all timings were the same. Observers were required to report the side on which the probe stimulus appeared to have the greater curvature in a 2AFC task. The repulsive nature of the tilt aftereffect results in reduced apparent curvature of both probes and any additional "curvature aftereffect" (CAE) would result in a further reduction. The staircase increased or decreased the contour angle of the test probe to home in on the point at which observers perceived the two probe stimuli to have equal curvature (PSE).

As in the plaid adaptation, a minimum of four staircases with the compound adaptor in the left hemi-field and four staircases with the compound adaptor in the right hemi-field were collected for each probe type and stimulus condition.

Data analysis

Each participant collected at least 200 trials for each condition with the compound adaptor on each side of fixation (4×50-trial staircases). The responses for each probe stimulus intensity level (either contrast difference or angular difference) were averaged for each observer. Depending on the adaptation task, either a logistic or Weibull function was fit to the averaged data. The PSE was derived from this fit as the point at which the observer was at 50% probability of responding on the compound side. Using this method, all data contribute to the calculation of the PSE, rather than only the trials on which reversals occur, and a full psychometric function can be recovered. It should be noted that data points near the PSE have more trials contributing to each point, as a result of the staircase procedure itself. Figure 2 shows a sample pair of psychometric curves.

For each condition, we quantified the magnitude of compound adaptation as the amount of additional contrast (plaid experiments) or curvature (contour experiments) required in the compound adapted hemi-field for the probes to appear equal. This is the mean shift in the PSE from the point of veridical equality. These average PSE values (selective adaptation) are plotted as differential effects. That is, selective adaptation effects are plotted as adaptation to the compound above and beyond adaptation to the components. Additionally, plaid adaptation effects are expressed in decibels using the following equation:

$$\mathrm{dB} = 20^{*}\log 10 \left(C_{\mathrm{adapt}}/C \right),$$

where C is the contrast of the reference probe and C_{adapt} is the Michelson contrast value required to equate the test probe to the reference.

Functions were fit to 5000 within-subject bootstrap re-samples for each condition (each with 200 trials for each side as in the original data set) so that, for each re-sample, a whole new pair of psychometric functions could be derived (one for each side on which the compound was adapted). The PSE values for each pair of functions were averaged to account for side bias then used to derive 95% confidence intervals of the PSE for each observer in each condition.

Experiment 1: Spatial frequency tuning

One way to ameliorate the combinatorial explosion in combining V1 neuron outputs is that neurons at the next stage ("conjunction detectors") could be invariant to certain features of the stimuli, such as spatial frequency. For example, a "curvature detector" might respond to a particular degree of curvature irrespective of the spatial frequency of the gratings from which it is constructed. By reducing the number of parameters being encoded, fewer processing units may be needed. In order to test this idea, Experiment 1 examined the spatial frequency tuning of the plaid adaptation effect and the contour adaptation effect, using an "adapt-one, test-many" design.

Experiment 1a: Plaid adaptation

Observers were adapted to stimuli with a spatial frequency (SF) of 1.26 c/deg. The SF of the probe stimuli took one of five values ranging from 0.4 to 4.0 c/deg depending on the trial. The different probe SFs were tested in separate sessions. One observer (DM) was tested on two other adapting SFs, 0.4 and 4.0 c/deg, for the full range of 5 probe SFs.

Figure 3A shows spatial frequency tuning curves for one participant (DM) collected for three different adapting spatial frequencies. Although the overall magnitude of adaptation varies with different adaptor spatial frequencies, plaid adaptation is always most apparent

when the adaptor and probes share a common spatial frequency. Figure 3B shows individual and group data for four participants collected with adapting stimuli of 1.26 c/deg. Again, peak adaptation effects of almost 3 dB are observed in trials where adaptor and probe share a common spatial frequency, with adaptation dropping to 1 dB for probes at either extreme of the function. These results indicate that the mechanisms responsible for processing plaid patterns are clearly tuned for spatial frequency. A Gaussian function was fitted to the group data and the tuning bandwidth (full-width half-height, FWHH) was estimated to be 2.72 octaves. These results are plainly at odds with the notion of spatial frequency-invariant conjunction detectors.

Experiment 1b: Contour adaptation

The spatial frequency of the adapting stimuli was fixed at 1.1 c/deg and the spatial frequency of the probe stimuli could take one of seven values between 0.4 and 3 c/deg (0.4, 0.56, 0.78, 1.1, 1.53, 2.14, 3.0 c/deg). The different probe SFs were tested in separate sessions. Five observers completed a minimum of four staircases for each probe SF on each side. One observer (SH) was tested on two other adapting SFs: 0.56 c/deg and 2.14 c/deg, for the full range of 7 probe SFs.

The magnitude of the contour adaptation effect, both for individual observers and averaged across observers, when adapting to 1.1 c/deg stimuli and testing with SFs between 0.4 and 3 c/deg is shown in Figure 4B. A repeated measures ANOVA was performed on the group data finding that the spatial frequency of the probe had a significant effect on the curvature aftereffect ($F(5.06, 20.23) = 3.756$, $p < 0.01$ with Huyn–Feldt correction). As for the plaid adaptation, the contour adaptation effect was greatest when the spatial frequencies of the adaptor and test stimuli were similar. When both adaptor and test stimuli had an SF of 1.1 c/deg, the effect magnitude was 4.76°. This dropped to 1.62° for a probe SF of 0.4 c/deg, and 1.3° for a probe SF of 3.0 c/deg. Paired t-tests demonstrated a significant difference in the magnitude of the CAE for probe SFs of 0.4 and 1.1 ($t(4) = -3.13$, $p < 0.05$), indicating that the CAE was reduced for probe SFs lower than the adapting frequency. Although the mean magnitude of the CAE at the highest probe SF (3.0 c/deg) was substantially lower than that when the adapting and probe SFs were equal, this reduction did not quite reach significance ($t(4) = 2.552$, $p = 0.063$) suggesting greater inter-observer variability at the highest probe SF. Indeed, some of the observers reported that the task was particularly difficult in this condition. Overall, there was little, if any, aftereffect for the lowest and highest SF probes, with 95% confidence intervals either included zero (highest SF probe) or had a lower bound very close to zero (lowest SF probe). A Gaussian function was fitted to these data and the tuning bandwidth (FWHH) was found to be 2.14 octaves.

Figure 4A shows the SF tuning of the contour adaptation effect for a single observer for three different adaptor SFs. Similar results are seen for all three adapting SFs. The greatest effect magnitude was found when the probe and adaptor SFs were identical. A bootstrapped re-sampling analysis of these data revealed that, for all three adapting SFs, there was a significant drop off in the magnitude of the curvature aftereffect as the probe SF became more different from that of the adaptor (the 95% confidence intervals for probe SFs of 0.4 c/deg and 3.0 c/deg did not overlap with those for the probe SF identical to the adapting SF). In fact, for the maximum adaptor/probe SF differences for adaptor SFs of 0.56 and 2.14 c/deg, no significant contour-selective adaptation effects were found.

These results clearly show that the curvature aftereffect is tuned for spatial frequency.

Experiment 2: Selective combinations of spatial frequencies

A second way to reduce the combinatorial explosion is for conjunction detectors only to combine representations of stimuli that share certain features. This may have further implications with respect to the perception of similar features as coming from a single coherent object. In Experiment 2 we examined the effect of varying the spatial frequency of the components comprising the plaid or the contour to see whether the plaid aftereffect and the curvature aftereffect are selective for the spatial frequencies within the adapting stimulus. Here the spatial frequency of each component was the same in the probe as it was in the adaptor, but the two components could differ relative to each other.

Experiment 2a: Plaid adaptation

The SFs of the component gratings within both the adaptors and the probe stimuli were varied across trials so that they were comprised either of two low SFs (0.4 c/deg, 0.4 c/deg), two mid-value SFs (1.6 c/deg, 1.6 c/deg), two high SFs (3.2 c/deg, 3.2 c/deg) or were separated by 1 octave (0.4 c/deg, 1.6 c/deg) or 3 octaves (0.4 c/deg, 3.2 c/deg). At a later date, the two participants who were still available for testing, out of the original four, collected data on two further conditions, one where both gratings had an SF of 6.4 c/deg and one where the two gratings had SFs separated by 4 octaves (0.4 c/deg, 6.4 c/deg). As in Experiment 1, the different SF conditions were tested in separate sessions.

Mean PSE shifts for individual participants and across the group are shown in Figure 5. For both individual and mean data the greatest adaptation was found when the plaid was comprised of components with a common spatial frequency. Although a moderate decrease is observed in the mean data when both component gratings had the same high spatial frequency, this decrease is within the margin of error. When the plaid components are separated by a single octave a clear decrease in plaid adaptation is observed, with a further decrease for a 3- or 4-octave separation. Thus, not all pairwise combinations of gratings lead to large compound adaptation effects. When the component gratings that comprise the plaid are substantially different, a reduced adaptive effect is observed. The diminished aftereffect observed in conditions with dissimilar components cannot be attributed to adaptation to a single grating; any of the SFs at which we tested resulted in a strong adaptation effect if both components had this SF. Therefore it seems that plaid adaptation is greatest when the gratings comprising the plaid have the same SF and decreases with increasing difference between the two gratings.

Experiment 2b: Contour adaptation

The spatial frequency of one component was fixed at 1.1 c/deg and the spatial frequency of the other component could take one of seven different values between 0.4 and 3 c/deg (0.4, 0.56, 0.78, 1.1, 1.53, 2.14, 3.0 c/deg). The different component SF conditions were tested in separate sessions. Three observers completed a minimum of four staircases for each component SF on each side.

The magnitude of the curvature aftereffect (PSE shift) for individual participants and averaged across the group is shown in Figure 6. The greatest adaptation effect was found when the two gratings comprising the contour had the same spatial frequency (mean = 5.38°). All observers showed a significant reduction in the CAE when the components differed substantially and showed no significant CAE for the greatest difference in SF between the component gratings (95% CIs included zero).

It should be noted that varying the spatial frequency of the gratings comprising the contour also disrupts the phase alignment between them. This could potentially be the reason for the observed reduction in the CAE, rather than the differing spatial frequencies per se. In order

to rule out this possibility, we conducted a control experiment for one observer (SH) where the two gratings had the same spatial frequency but were either phase aligned or 180° degrees out of phase. Mean PSE shifts in the aligned and 180° out of phase conditions were 3.92° and 4.52°, respectively, which were not significantly different (95% confidence intervals overlapped). Furthermore, the results for the out of phase condition was not significantly different from the data for that observer in the same SF condition in the main experiment, suggesting that phase misalignment is not the reason for the disruption in contour-selective adaptation with different spatial frequency components.

Thus, the curvature aftereffect is also clearly tuned for the spatial frequency of individual component gratings.

Discussion

It is well known that the receptive fields of V1 neurons respond well to grating stimuli of specific orientations and spatial frequencies. How the outputs of these cells are combined at subsequent stages in visual processing is less well known. Stimulus-selective aftereffects suggest that the outputs of neurons with overlapping and adjacent receptive fields can be combined to represent plaids and contours, respectively (Hancock & Peirce, 2008; Peirce & Taylor, 2006). However, it is a well-known problem that the huge number of possible combinations makes it highly improbable that all pairwise combinations of V1 neuron outputs are represented. We present two possible solutions to the problem. One is that these putative conjunction detectors for plaids and contours may be invariant to some lower level features of the previous representation, for example spatial frequency, allowing many combinations of V1 outputs to be represented by a single detector. Alternatively, these detectors could employ a strategy of selective combination, whereby only those combinations of outputs that share similar features are represented. The current study examined the spatial frequency tuning of the plaid contrast aftereffect and the contour shape aftereffect to investigate how V1 outputs are combined in order to ameliorate this "combinatorial explosion."

Spatial frequency tuning

When the spatial frequency of the test stimulus was varied relative to the adaptor, both the plaid and the curvature aftereffects demonstrated a clear tuning for spatial frequency. The greatest aftereffects occurred when the test pattern and the adaptor had the same spatial frequency. Thus, it appears that the spatial frequency information in the stimulus is retained by the conjunction detector and, therefore, that multiple detectors must be needed to represent the range of frequencies.

The reduced effects do not result simply from a reduced sensitivity to high spatial frequencies, due to the stimuli being presented in peripheral locations. When both adaptor and probe stimuli have high spatial frequencies the magnitude of adaptation is comparable to other conditions (see Figures 3A and 4A). It is only when the spatial frequencies differ between adaptor and probe or between components that the effects weaken.

The spatial frequency tuning for the CAE (2.14 octaves at full-width half-height; FWHH) is consistent with spatial frequency tuning of other low- and mid-level shape aftereffects such as the tilt aftereffect (TAE). For example, Ware and Mitchell (1974) report a half-width half-height bandwidth of 1–2 octaves, equivalent to a bandwidth of 2–4 octaves at full-width half-height as calculated here. Similarly, the shape frequency aftereffect (SFAE, Gheorghiu & Kingdom, 2006) was reduced when the adapting and test stimuli differed in "luminance scale" (coarse versus fine), which is essentially the spatial frequency of the carrier forming the contour. Gheorghiu and Kingdom also found a degree of selectivity to both luminance

polarity and even a small preference for equal contrast between test and adapting stimuli. They concluded that contour shape was encoded in a relatively "feature-rich" form, retaining information about these low-level features.

A number of studies have looked at the effects of spatial frequency on contrast adaptation with gratings, with somewhat mixed results. Contrast threshold elevation aftereffects are clearly tuned for spatial frequency with (FWHH) bandwidth estimates of around 1–2 octaves (Blakemore & Campbell, 1969; Snowden & Hammett, 1996). When the test stimulus is presented at a supra-threshold level, some authors have found the reduction in perceived contrast after adaptation to be largely untuned for spatial frequency (Snowden & Hammett, 1996), whereas other authors have found clear tuning (Blakemore, Muncey & Ridley, 1973). Here, we report spatial frequency tuning of plaid contrast adaptation, with an estimated bandwidth of 2.72 octaves, consistent with the underlying mechanisms being band-pass tuned for spatial frequency.

This spatial frequency tuning indicates that conjunction mechanisms do not pool across all spatial frequency channels in combining local orientation signals from V1. This, apparently, is not the approach the visual system uses to form an economical representation of information beyond V1 suggesting that it must employ an alternative strategy to reduce the number of combinations of outputs that are represented.

Selective combinations of spatial frequencies

When the two component gratings comprising the adapting plaids or contours had different spatial frequencies, we again found that both the plaid and curvature aftereffects were selective for spatial frequency. That is, both aftereffects were maximal when the two gratings had the same spatial frequency, with adaptation effects clearly reduced when the gratings comprising the compound were different. This suggests that the less similar the low-level features, the less likely their representative outputs are to be combined. The tuning bandwidth (FWHH) of the curvature effect with respect to the relative spatial frequencies of the components was 2.3 octaves, very much in keeping with the bandwidths previously measured for lower level mechanisms such as the TAE (Ware & Mitchell, 1974), presumably reflecting the tuning of the input filters. For the plaid aftereffect, the tuning seems broader at roughly 6 octaves (this can only be estimated because only the high-SF flank of the function is available).

Again, this spatial frequency selectivity in the curvature aftereffect is consistent with previous findings of band-pass spatial frequency tuning in other tasks that involve integration of local components. Path detection in Gabor field paradigms (Dakin & Hess, 1998), detecting spatial jitter in circular arrays of Gabors (Keeble & Hess, 1999), local (within dipole) integration in glass patterns, and coherence thresholds for local integration in random dot motion (Bex & Dakin, 2002) are all disrupted by varying the spatial frequency of individual components.

One of the implications of the spatial frequency selectivity of these aftereffects is that components are only combined when they share broadly similar features. As a result these mechanisms may be useful in the Gestalt grouping (or "perceptual binding") of similar visual components. For drifting plaid patterns, it is well established that when the gratings comprising the plaid have similar spatial frequencies (within 1.5–2 octaves) they are perceived as a single coherent pattern, whereas for substantially different components the percept is of two semitransparent gratings sliding over each other (Adelson & Movshon, 1982; Kim & Wilson, 1993; Smith, 1992; Stoner & Albright, 1992). The exact spatial frequency limits for this coherence also depend on the angular difference between the component directions (Kim & Wilson, 1993) and on the contrast of the pattern (Smith,

1992). Huk and Heeger (2002) used fMRI adaptation to demonstrate that percepts of coherent motion were closely linked to the activity of pattern motion-selective neurons in human MT+. Reducing the perceptual coherence of pattern motion by adjusting the spatial frequency of the component gratings to differ by 3 octaves produced a corresponding decrease in adaptation to plaid motion in MT but not V1. In the current study, the plaids were not drifting and so percepts of transparency per se do not occur. Nonetheless, participants report that they can identify the separate component gratings of the plaids with dissimilar components more readily than those with identical components. This is consistent with reports by Georgeson (1998) that plaids were perceived as two separate components when their spatial frequencies differed by up to 1.5 octaves. For such stimuli in our study, the aftereffect was dramatically reduced. Thus, these simple conjunction detectors may have a role in the perceptual binding of gratings as coherent plaids and as continuous contours.

The selectivity of the plaid adaptation effect with respect to spatial frequency is also very much in keeping with the notion of the "doughnut" mechanisms of Olzak and Thomas (1999), which combine gratings over a wide range of orientations but a narrow range of spatial frequencies. Nam, Solomon, Morgan, Wright, and Chubb (2009) have also recently shown pop-out effects for plaids among a field of gratings but only when both plaid components were the same spatial frequency.

Processing at higher levels

While it appears that low and intermediate levels of visual processing rely on feature-selective binding of earlier signals to increase coding efficiency, the same may not be true at higher levels in visual processing. Differences between very similar low- and high-level aftereffects reveal a dichotomy in spatial frequency tuning between feature and global form processing. Both the current data and the shape amplitude aftereffect described by Gheorghiu and Kingdom (2006) show spatial frequency tuning in the processing of curved contours. However, when the contour is closed to form a radial frequency pattern the observed amplitude aftereffect (RFAAE) is not tuned for spatial frequency (Bell & Kingdom, 2009) and detection thresholds for radial frequency patterns are similarly untuned (Wilkinson, Wilson, & Habak, 1998). A recently described higher level form of the tilt after-effect, generated by spatially remote adaptation to a global figure, was also reported to be broadband tuned for spatial frequency (Roach, Webb, & McGraw, 2008). Other processes that involve global integration, such as sensitivity to overall structure in glass patterns (Dakin & Bex, 2001) and Gabor arrays (Achtman, Hess, & Wang, 2003) and to random dot motion (Bex & Dakin, 2002) all appear to be more broadly tuned, if at all, for spatial frequency.

The compound adaptation effects used here appear to probe intermediate mechanisms, different to both the low-level contrast adaptation of orientation detectors, but also dissimilar, at least in spatial frequency tuning, to the global form mechanisms described by numerous other authors.

Summary

Although previous work has shown that spatial frequency channels can be combined, particularly for the detection of edges or lines (e.g., Marr & Hildreth, 1980), the results presented here suggest that not all mid-level mechanisms combine across spatial frequency in this way. Our results show that the mechanisms underlying both plaid adaptation and contour adaptation aftereffects are selective for the spatial frequency of the component gratings of which the plaid or contour is comprised. Rather than conjunction detectors being invariant to low-level features, one way in which the combinatorial explosion might be dealt with in the visual system is by selectively combining only those input channels with similar

features, dramatically reducing the number of combinations that are encoded. The fact that this is observed in two qualitatively different aftereffects may suggest that it is a general feature of mid-level visual processing. One result of this selective combining is that these mechanisms may be used in the Gestalt grouping or "binding" of visual features. Co-first authors DPM and SH contributed equally to this research.

Acknowledgments

References

Achtman RL, Hess RF, Wang Y-Z. Sensitivity for global shape detection. Journal of Vision. 2003; 3(10):616–624. 4. http://www.journalofvision.org/content/3/10/4, doi:10.1167/3.10.4. [PubMed: 14640885]

Adelson EH, Movshon JA. Phenomenal coherence of moving visual patterns. Nature. 1982; 300:523–525. [PubMed: 7144903]

Anzai A, Peng X, Van Essen DC. Neurons in monkey visual area V2 encode combinations of orientations. Nature Neuroscience. 2007; 10:1313–1321.

Bell J, Kingdom FAA. Global contour shapes are coded differently from their local components. Vision Research. 2009; 49:1702–1710. [PubMed: 19375448]

Bex PJ, Dakin SC. Comparison of the spatial-frequency selectivity of local and global motion detectors. Journal of the Optical Society of America A. 2002; 19:670–677.

Blakemore C, Campbell FW. On the existence of neurons in the human visual system selectively sensitive to the orientation and size of retinal images. The Journal of Physiology. 1969; 203:237–260. [PubMed: 5821879]

Blakemore C, Muncey JPJ, Ridley RM. Stimulus specificity in the human visual system. Vision Research. 1973; 13:1915–1931. [PubMed: 4746989]

Dakin SC, Bex PJ. Local and global visual grouping: Tuning for spatial frequency and contrast. Journal of Vision. 2001; 1(2):99–111. 4. http://www.journalofvision.org/content/1/2/4, doi: 10.1167/1.2.4. [PubMed: 12678605]

Dakin SC, Hess RF. Spatial-frequency tuning of visual contour integration. Journal of the Optical Society of America A. 1998; 15:1486–1499.

Felleman DJ, Van Essen DC. Distributed hierarchical processing in the primate cerebral cortex. Cerebral Cortex. 1991; 1:1–47. [PubMed: 1822724]

Gallant JL, Braun J, Van Essen DC. Selectivity for polar, hyperbolic and Cartesian gratings in Macaque visual cortex. Science. 1993; 259:100–103. [PubMed: 8418487]

Gallant JL, Connor CE, Rakshit S, Lewis JW, Van Essen DC. Neural responses to polar, hyperbolic and Cartesian gratings in area V4 of the Macaque monkey. Journal of Neurophysiology. 1996; 76:2718–2739. [PubMed: 8899641]

Georgeson MA. Edge-finding in human vision: A multi-stage model based on the perceived structure of plaids. Image and Vision Computing. 1998; 16:389–405.

Georgeson MA, Meese TS. Perception of stationary plaids: The role of spatial filters in edge analysis. Vision Research. 1997; 37:3255–3271. [PubMed: 9425542]

Gheorghiu E, Kingdom FAA. Luminance-contrast properties of contour-shape processing revealed through the shape-frequency aftereffect. Vision Research. 2006; 46:3603–3615. [PubMed: 16769101]

Gheorghiu E, Kingdom FAA. The spatial feature underlying the shape-frequency and shape-amplitude after-effects. Vision Research. 2007; 47:834–844. [PubMed: 17292437]

Hancock S, Peirce JW. Selective mechanisms for simple contours revealed by compound adaptation. Journal of Vision. 2008; 8(7):1–10. 11. http://www.journalofvision.org/content/8/7/11, doi: 10.1167/ 8.7.11.

Henning GB, Hertz BG, Broadbent DE. Some experiments bearing on the hypothesis that the visual system analyses spatial patterns in independent bands of spatial frequency. Vision Research. 1975; 15:887–897. [PubMed: 1166622]

Huk AC, Heeger DJ. Pattern motion responses in human visual cortex. Nature Neuroscience. 2002; 5:72–75.

Keeble DRT, Hess RF. Discriminating local continuity in curved figures. Vision Research. 1999; 39:3287–3299. [PubMed: 10615496]

Kim J, Wilson HR. Dependence of plaid motion coherence on component grating directions. Vision Research. 1993; 33:2479–2489. [PubMed: 8249328]

Ledgeway T, Smith AT. Evidence for separate motion-detecting mechanisms for first- and second-order motion in human vision. Vision Research. 1994; 34:2727–2740. [PubMed: 7975310]

Marr D, Hildreth E. Theory of edge detection. Proceedings of the Royal Society of London B. 1980; 207:187–217.

McGovern D, Peirce JW. The spatial characteristics of plaid-form-selective mechanisms. Vision Research. 2010; 50:796–804. [PubMed: 20123109]

Morgan MJ, Watt RJ. The combination of filters in early spatial vision: A retrospective analysis of the MIRAGE model. Perception. 1997; 26:1073–1088. [PubMed: 9509144]

Morrone MC, Burr DC. Feature detection in human vision: A phase-dependent energy model. Proceedings of the Royal Society of London B. 1988; 235:221–245.

Movshon, JA.; Adelson, EH.; Gizzi, MS.; Newsome, WT. The analysis of moving patterns. In: Chagas, C.; Gattass, R.; Gross, C., editors. Pattern recognition mechanisms. Springer; New York: 1985. p. 117-151.

Nam J-H, Solomon J, Morgan MJ, Wright CE, Chubb C. Coherent plaids are preattentively more than the sum of their parts. Attention, Perception & Psychophysics. 2009; 71:1469–1477.

Olzak LA, Thomas JP. When orthogonal orientations are not processed independently. Vision Research. 1991; 31:51–57. [PubMed: 2006554]

Olzak LA, Thomas JP. Configural effects constrain Fourier models of pattern discrimination. Vision Research. 1992; 32:1885–1898. [PubMed: 1287986]

Olzak LA, Thomas JP. Neural recoding in human pattern vision: Model and mechanisms. Vision Research. 1999; 39:231–256. [PubMed: 10326133]

Pasupathy A, Connor CE. Responses to contour features in macaque area V4. Journal of Neurophysiology. 1999; 82:2490–2502. [PubMed: 10561421]

Pasupathy A, Connor CE. Shape representation in area V4: Position-specific tuning for boundary conformation. Journal of Neurophysiology. 2001; 86:2505–2519. [PubMed: 11698538]

Pasupathy A, Connor CE. Population coding of shape in area V4. Nature Neuroscience. 2002; 5:1332–1338.

Peirce JW. The potential importance of saturating and supersaturating contrast response functions in visual cortex. Journal of Vision. 2007a; 7(6):1–10. 13. http://www.journalofvision.org/content/7/6/13, doi:10.1167/7.6.13.

Peirce JW. PsychoPy—psychophysics software in Python. Journal of Neuroscience Methods. 2007b; 162:8–13. [PubMed: 17254636]

Peirce JW, Taylor LJ. Selective mechanisms for complex visual patterns revealed by adaptation. Neuroscience. 2006; 141:15–18. [PubMed: 16753271]

Roach NW, Webb BS, McGraw PV. Adaptation to global structure induces spatially remote distortions of perceived orientation. Journal of Vision. 2008; 8(3):1–12. 31. http://www.journalofvision.org/content/8/3/31, doi:10.1167/8.3.31.

Rust NC, Mante V, Simoncelli EP, Movshon JA. How MT cells analyze the motion of visual patterns. Nature Neuroscience. 2006; 9:1421–1431.

Smith AT. Coherence of plaids comprising components of disparate spatial frequencies. Vision Research. 1992; 32:393–397. [PubMed: 1574854]

Snowden RJ, Hammett ST. Spatial frequency adaptation: Threshold elevation and perceived contrast. Vision Research. 1996; 36:1797–1809. [PubMed: 8759448]

Stoner GR, Albright TD. Motion coherency rules are form-cue invariant. Vision Research. 1992; 32:465–475. [PubMed: 1604833]

Ware C, Mitchell DE. The spatial selectivity of the tilt aftereffect. Vision Research. 1974; 14:735–737. [PubMed: 4424240]

Watt RJ, Morgan MJ. A theory of the primitive spatial code in human vision. Vision Research. 1985; 25:1661–1674. [PubMed: 3832590]

Wilkinson F, Wilson HR, Habak C. Detection and recognition of radial frequency patterns. Vision Research. 1998; 38:3555–2947. [PubMed: 9893789]
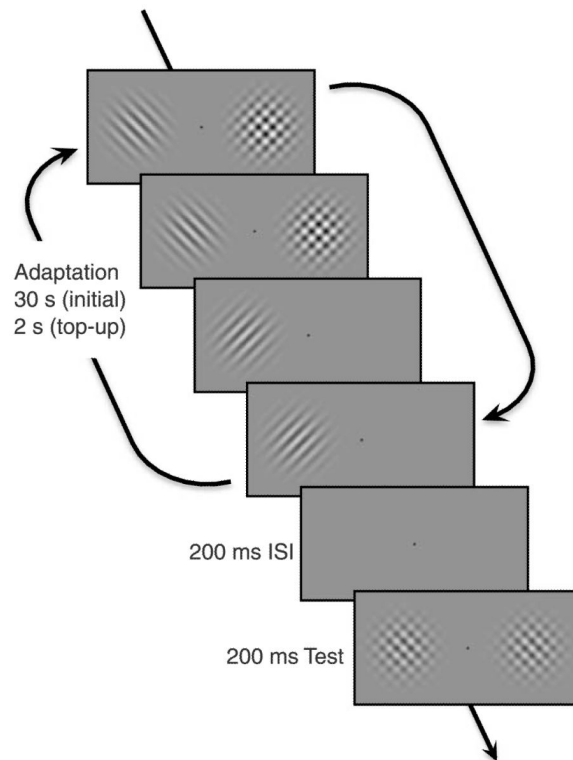
Figure 1.
Schematic of the procedure for Experiment 1. In the component field (left), individual gratings alternated every second, and in the compound field (right), the compound stimulus alternated with a blank field every second. Note that for convenience these are depicted as alternating at the same times. In the actual experiment, the relative temporal phase of the alternations in the two hemi-fields was given a random offset that differed every trial. For this experiment, the reference probe contrast was fixed at 42% and the test probe contrast (compound side) was varied according to a staircase procedure to find the PSE. The probe stimuli consisted of compound stimuli with the same orientation as the adapting stimuli. On each trial, observers indicated which probe stimulus appeared to have the higher contrast. Each observer performed a minimum of four staircases for each adapting configuration.
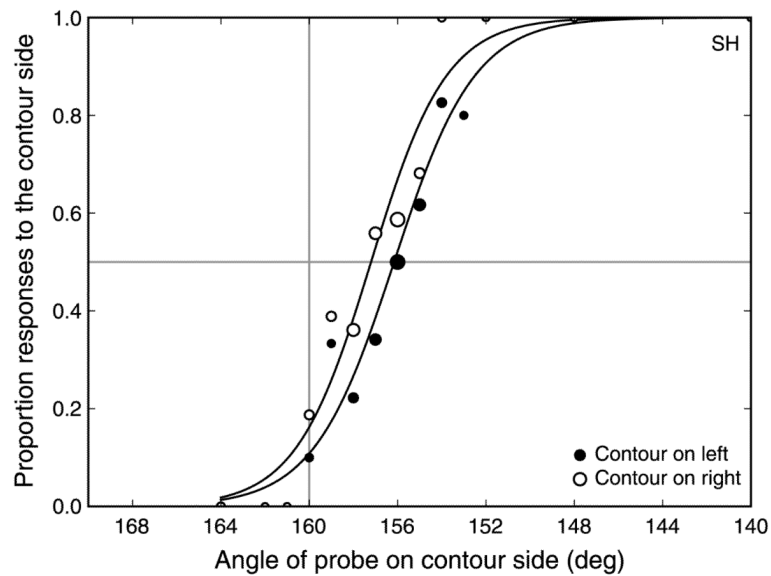
Figure 2.
Sample psychometric functions for compound stimuli on the left (filled symbols) and right
(open symbols) of fixation (Experiment 1b, observer SH, adapt SF = 1.1 c/deg, test SF =
0.78 c/deg). Symbol size indicates the relative number of responses underlying each data
point (actual numbers vary from 1 to 70 responses per point). The horizontal gray line
indicates the point of 50% probability of responding on the compound side. The vertical
gray line (160° on *x*-axis) represents the point of veridical equality. The PSE is calculated
from the average of the points where the left and right side functions cross the 50%
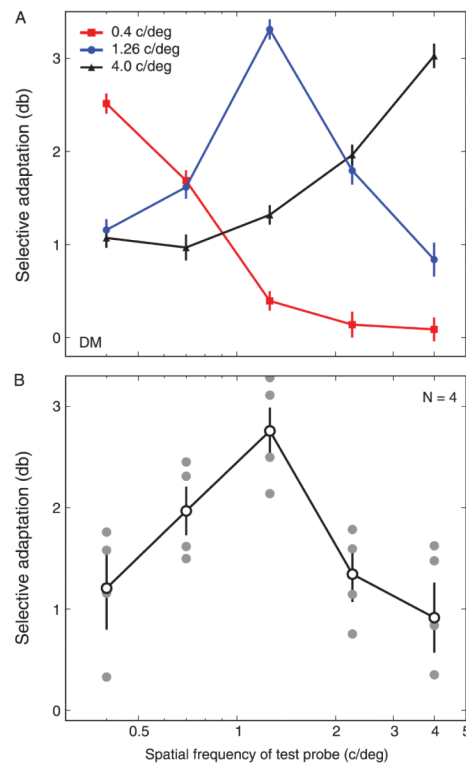probability line.

Figure 3.
Magnitude of the plaid adaptation effect for different adapting and test spatial frequencies. (A) Data from observer DM for adapting SFs of 0.4, 1.26, and 4.0 c/deg. Error bars indicate 95% CIs based on 5000 bootstrap re-samples. (B) Individual subject data (filled symbols) and group mean data (open symbols, $N = 4$) adapting to 1.26 c/deg and testing with SFs between 0.4 and 4 c/deg. Error bars indicate *SEM* across observers; 0 on the *y*-axis indicates the point of veridical equality.
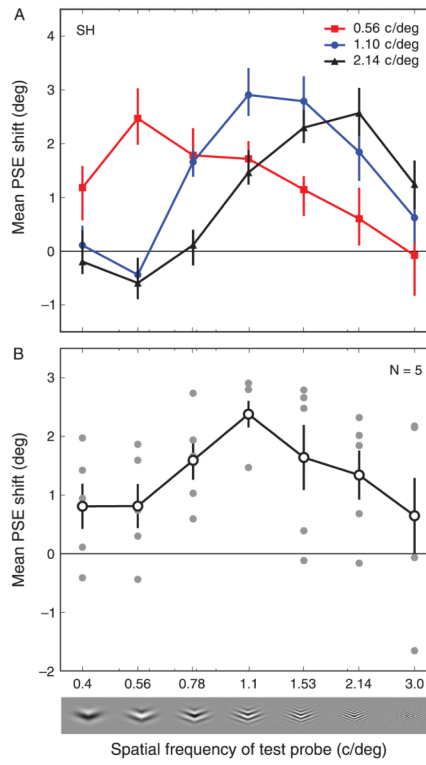
Figure 4.
Magnitude of the contour adaptation effect for different adapting and test spatial frequencies. (A) Data from observer SH for adapting SFs of 0.56, 1.1, and 2.14 c/deg. Error bars indicate 95% CIs based on 5000 bootstrap re-samples. (B) Individual subject data (filled symbols) and group mean data (open symbols, $N = 5$) adapting to 1.1 c/deg and testing with SFs between 0.4 and 3 c/deg. Error bars indicate *SEM* across observers; 0 on the *y*-axis indicates the point of veridical equality.
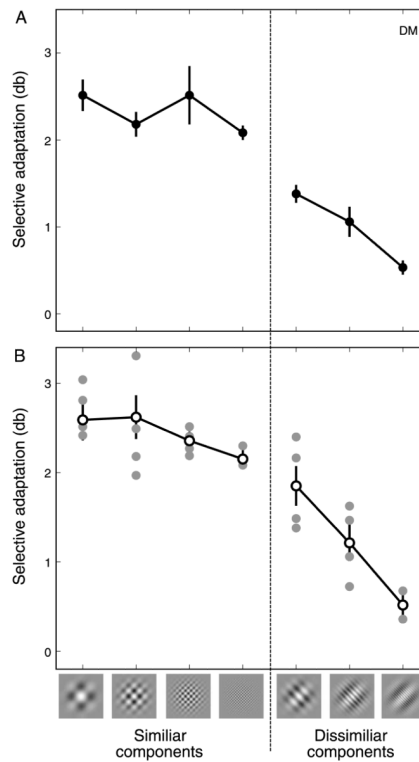
Figure 5.
Magnitude of the plaid adaptation effect for plaids comprised of component gratings with the same or different spatial frequencies. (A) Data from observer DM. Error bars indicate 95% CIs based on 5000 bootstrap re-samples. (B) Individual subject data (filled symbols) and group mean data (open symbols). For the group mean data, the rightmost data points for the similar and dissimilar components represent data from two observers (see General methods section for details). Otherwise, $N = 4$. Error bars indicate *SEM* across observers; 0 on the *y*-axis indicates the point of veridical equality.
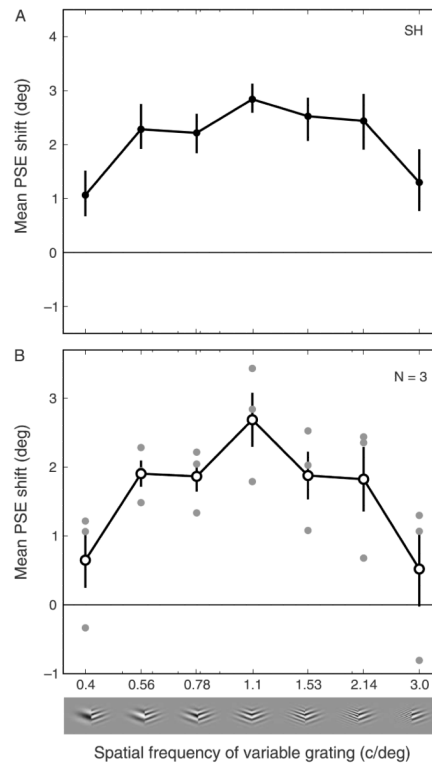
Figure 6.
Magnitude of the curvature adaptation effect for contours comprised of component gratings with the same or different spatial frequencies. (A) Data from observer SH. Error bars indicate 95% CIs based on 5000 bootstrap re-samples. (B) Individual subject data (filled symbols) and group mean data (open symbols, $N = 3$). Error bars indicate *SEM* across observers; 0 on the *y*-axis indicates the point of veridical equality.