Pattern masking: The importance of remote spatial frequencies and their phase alignment

Pi-Chun Huang
McGill Vision Research, Department of Ophthalmology, Montreal, Quebec, Canada

Goro Maehara
Graduate School of Arts and Sciences, University of Tokyo, Tokyo, Japan

Keith A. May
Department of Optometry and Visual Science, City University, London, UK

Robert F. Hess
McGill Vision Research, Department of Ophthalmology, Montreal, Quebec, Canada

To assess the effects of spatial frequency and phase alignment of mask components in pattern masking, target threshold vs. mask contrast (TvC) functions for a sine-wave grating (S) target were measured for five types of mask: a sine-wave grating (S), a square-wave grating (Q), a missing fundamental square-wave grating (M), harmonic complexes consisting of phase-scrambled harmonics of a square wave (Qp), and harmonic complexes consisting of phase-scrambled harmonics of a missing fundamental square wave (Mp). Target and masks had the same fundamental frequency (0.46 cpd) and the target was added in phase with the fundamental frequency component of the mask. Under monocular viewing conditions, the strength of masking depends on phase relationships among mask spatial frequencies far removed from that of the target, at least 3 times the target frequency, only when there are common target and mask spatial frequencies. Under dichoptic viewing conditions, S and Q masks produced similar masking to each other and the phase-scrambled masks (Qp and Mp) produced less masking. The results suggest that pattern masking is spatial frequency broadband in nature and sensitive to the phase alignments of spatial components.

Keywords: pattern masking, spatial frequency, phase alignment, masking, dichoptic blur suppression


Introduction

Our knowledge of spatial processing has benefited from the numerous studies carried out on pattern masking. Typically, pattern masking is studied by evaluating the detectability of a target in the presence of masks of different contrasts (TvC function). When the target and mask stimuli are identical, the task amounts to contrast discrimination. In this case, the TvC function has a characteristic non-monotonic form, called a dipper function. A dip occurs at low mask contrasts, where low-contrast masks facilitate the detection of the target, and a rising limb with slope around 0.6 occurs at high mask contrasts, where higher contrast masks suppress the detection of the target (Legge, 1981; Legge & Foley, 1980). The dipper function is commonly explained as the consequence of a contrast transducer (Foley & Legge, 1981; Legge & Foley, 1980; Nachmias & Sansbury, 1974), contrast gain control (Foley, 1994; Foley & Chen, 1999), or stimulus uncertainty (Klein & Levi, 2009; Pelli, 1985). When mask and target are presented to different eyes (dichoptic viewing condition), the dipper effect is reduced and more pronounced masking effects are found with the slope of the rising limb approaching unity, i.e., Weber’s law (Legge, 1979; Maehara & Goryo, 2005; Meese, Georgeson, & Baker, 2006). The dichoptic TvC function can also be explained by a contrast gain control process before or after binocular interactions (Baker, Meese, & Summers, 2007; Maehara & Goryo, 2005; Meese et al., 2006).

Although these explanations have been successful in describing the results for narrowband stimuli for monocular and dichoptic viewing conditions, they fail when broadband stimuli (e.g., noise) are involved in pattern masking. For example, pedestal effects that persist for additional broadband noise masks disappeared in notched noise masks (Henning & Wichmann, 2007). Following this research, Goris, Wichmann, and Henning (2009) proposed a population code model where more than one detecting mechanism is involved. It has also been reported that masks with orientations different from target orientation could elevate thresholds (cross-orientation masking; Foley,
the sine-wave target and the missing fundamental mask. The square-wave mask includes “within-scale” and “cross-scale” masking effects; (2) whether or not masking remains constant if the mask’s contrast energy spectrum remains constant but relative phase relationships among higher harmonics change; and (3) the dependence of monocular/dichoptic viewing on the above two issues. Our results suggest that masking is spatial frequency broadband in nature and sensitive to the phase alignments of spatial components.

Methods

Observers

Three observers (GM, PCH, and TD) with corrected-to-normal vision participated in the experiment. PCH and GM are two of the authors.

Apparatus

Stimuli were generated using a VSG 2/5(Cambridge Research System, UK), which produces 15-bit gray-level resolution, and presented on a CRT monitor (SONY Triniton G520). The resolution of the monitor was 1024 × 768 pixels with refresh rate of 120 Hz. In order to manipulate the contrast of masks and targets independently, the target and mask were presented in alternate frames (interlaced technique). Therefore, the highest contrast was 50%. The mean luminance was 44.5 cd/m².

Stimuli

Five types of mask patterns were used: sine-wave grating (S), square wave (Q), missing fundamental square wave (M), harmonic complexes consisting of phase-scrambled harmonics of a square wave (Qp), and harmonic complexes consisting of phase-scrambled harmonics of a missing fundamental square wave (Mp; see the first row of Figure 1).

A sine-wave grating is defined as

$$S(x) = L_0[1 + c \sin(2\pi f(x - x_0))],$$  \hspace{1cm} (1)

where $L_0$ is mean luminance, $f$ is spatial frequency, $x_0$ is a random central location that shifts the phase of the sine-wave grating, and $c$ is the Michelson contrast defined as

$$c = \frac{c_{\text{max}} - c_{\text{min}}}{c_{\text{max}} + c_{\text{min}}}. \hspace{1cm} (2)$$
A square-wave grating can be viewed as the sum of an infinite series of sine-wave gratings:

\[
Q(x) = L_0 \left[ 1 + c \sum_{j=1}^{\infty} \frac{\sin((2j-1)2\pi f(x-x_0) + \theta_j)}{(2j-1)} \right]
\]

\[
= L_0 \left[ 1 + c \left( \sin(2\pi f(x-x_0) + \theta_1) \right.ight.
\]

\[
+ \left. \frac{1}{3} \sin(6\pi f(x-x_0) + \theta_2) \right]
\]

\[
+ \left. \frac{1}{5} \sin(10\pi f(x-x_0) + \theta_3) + \ldots \right] \right], \quad (3)
\]

where \( f \) is its fundamental frequency, \( c \) is the Michelson contrast of the fundamental spatial frequency of the square wave, and \( \theta_j \) is the phase of each component, \( j \). For a square wave (Q), the phase of each component is set to zero. For a Qp mask, \( \theta_1 \) is set to zero but the rest are randomized and a uniform probability distribution was used to assign the phases. The missing fundamental pattern (M) is the square wave minus its fundamental frequency component:

\[
M(x) = L_0 \left[ 1 + c \sum_{j=2}^{\infty} \frac{\sin((2j-1)2\pi f(x-x_0) + \theta_j)}{(2j-1)} \right]
\]

\[
= L_0 \left[ 1 + c \left( \frac{1}{3} \sin(6\pi f(x-x_0) + \theta_2) \right.ight.
\]

\[
+ \left. \frac{1}{5} \sin(10\pi f(x-x_0) + \theta_3) + \ldots \right] \right]. \quad (4)
\]

For the M configuration, all the \( \theta \) values are set to zero; for Mp configuration, the \( \theta \) values are randomized and \( c \) is the Michelson contrast of the fundamental spatial frequency of the corresponding square wave.

The images were patches of vertical gratings computed as described above with a central elliptic plateau of 5.34 \( \times \) 2.67 degrees (these are the lengths of the two main axes of the ellipse) and the contrast at the edge decayed according to a half-Gaussian function with standard deviation of 0.375 degree. The spatial frequency was 0.46 cpd with viewing distance of 114 cm. The stimuli were presented monocularly or dichoptically, which was achieved with a mirror stereoscope. We used a relatively low spatial frequency so that the harmonics would be easily detected and their influence, therefore, easier to gauge. The mask was one of the five gratings described above. The absolute phase of the target and mask was changed from trial to trial and the phase-scrambled masks were regenerated for each trial. The target stimulus was a sine-wave grating added in phase with the fundamental frequency component of the mask.

**Procedure**

A spatial three-alternative forced-choice, odd-man-out paradigm with unlimited viewing (determined by the subject’s response) was used in order to reduce the structure/feature influences on contrast discrimination task. This facilitated the detectability of stimuli whose spatial appearance can change with contrast, especially when the mask is M. Initially, we tried a temporal 2AFC task and subjects found that they could tell the difference between intervals but could not decide which interval contained the sine-wave target in the M mask configuration. This spatial task provided a sensitive method of making contrast discrimination.

![Figure 1. Illustration of the experimental stimuli and their profiles.](http://jov.arvojournals.org/pdfaccess.ashx?url=/data/journals/jov/932802/ on 06/20/2017)
comparisons without any time/feature constraints. A threedown–one-up staircase terminated at 6 upward reversals and the last five upward reversals were averaged to determine the threshold that is equal to 79% correct performance. Three staircase measures were collected to determine the averaged thresholds and standard errors for each condition. Three stimuli that had the same spatial properties were aligned vertically with separation of 3 degrees with each other (from center to center). While each stimulus contained a mask, a target was presented on one of them, chosen randomly. Observers had to indicate which one was different from the other two patches. In order to aid peripheral fusion and prevent binocular rivalry, the pairs of rectangular frames (6 degrees by 10 degrees) for each eye with one short line at the top for one rectangular frame and one short line at the bottom for the other rectangular frame were present throughout the whole experiment. Observers were pre-cued to trial onset by an audible tone and a fixation point; the latter consisted of an inverted T for one eye and upright T for the other eye, which were easily fused to form a cross. The fixation lines were presented for 200 ms and then disappeared. After a 200-ms delay, the stimuli were presented. To ensure that the participants had enough time to compare the 3 stimuli, the stimuli remained in view until observers made a response. No rivalry was seen in the dichoptic viewing condition between any of our stimulus configurations. The time to respond was typically 1–2 s. Feedback was given after each response. Seven levels of mask contrast were tested for each mask pattern and the order of the conditions (5 mask types × 7 mask contrasts) was randomized.

**Results**

**The effects of viewing condition**

Target contrast thresholds as a function of mask contrast for different masks and viewing configurations are plotted in Figure 2. In order to compare S and Q masks, results are plotted in terms of the contrast of the fundamental
sine-wave component. For M masks, the values on the x-axis represent the contrast of the missing fundamental frequency needed to form the Q mask.

The monocular masking results for the S mask (sine wave; green filled circles) exhibited a typical dipper shape on these logarithmic axes: The threshold decreased at low mask contrasts then increased linearly with increasing mask contrast. The dichoptic results for the S mask (green open circles) exhibit a reduced dipper effect and more pronounced masking effects that have been reported before (Legge, 1979; Maehara & Goryo, 2005; Meese et al., 2006). The slope at the rising limb is near unity (Weber’s law). For the Q and Qp masks (square wave; red and magenta squares), a similar pattern masking effect to that of the S mask was found under monocular and dichoptic viewing configurations. For the M mask (missing fundamental; blue filled diamonds), there was little or no facilitatory effect at low mask contrasts but masking effect at high mask contrasts under both viewing conditions and much stronger masking at higher mask contrasts under the dichoptic viewing configuration. For the Mp mask condition (cyan diamonds), both of the observers showed neither facilitation nor masking at low mask contrast but a masking effect at high mask contrasts under both viewing configurations, with the masking effect being the same for both viewing conditions.

In summary, dichoptic masking was stronger than monocular masking in all conditions except Mp, where the phase scrambling made the dichoptic mask no more effective than its monocular counterpart.

The effect of phase alignment

We replotted the results in Figure 3 to compare the masking effect between phase-aligned and phase-scrambled configurations. Generally speaking, there was a clear effect of phase alignment, both under monocular and dichoptic viewing conditions, with phase-aligned masks causing stronger masking than phase-scrambled masks even though they have the same contrast energy at each spatial frequency. This suggests the importance of phase alignment in pattern masking. The only exception is M and Mp monocular masks: Subject TD showed similar masking effect for M and Mp, although PCH showed comparable masking only at higher mask contrasts. The effect of phase alignment was more robust for the dichoptic conditions: Dichoptic masking results from the M and Mp masks, and from the Q and Qp masks, demonstrate that both cross- and within-scale dichoptic masking are dependent on the phase of mask frequencies far removed from that of the target.

The effect of higher harmonics

Figure 4 replots the results to show the effect of higher harmonics. Under the monocular viewing configuration,
the Q mask produced marginally stronger masking than the S mask at higher contrast levels and marginally weaker facilitation at low mask contrasts. For the M mask, there was a masking effect at high mask contrasts although it was not as strong as that for the Q mask. The Qp mask (magenta squares) produced a typical dipper function that was similar to the S mask. The Mp mask caused a masking effect at high mask contrasts but little (if any) facilitation. Thus, monocularly, stimuli with higher harmonics produced stronger masking than S, especially for within-scale stimuli (Q) whose components were phase aligned.

However, the phase of these additional frequency components was important, with phase-scrambled stimuli producing even less dichoptic masking, especially in the cross-scale case (Mp).

**Discussion**

We have three conclusions. The first conclusion concerns monocular versus dichoptic masking (Figure 2). Dichoptic masking is always stronger than monocular masking except in the case of pure cross-scale masking where the mask components are not phase aligned (Mp stimulus). In this case, monocular and dichoptic masking are similar. The second conclusion concerns the importance of the phase alignment of mask components (Figure 3). Phase-aligned mask components always produce stronger masking, except...
in the case of monocular cross-scale stimuli (M and Mp stimuli). The third conclusion concerns the importance of mask frequencies far removed from that of the target, at least 3 times the target frequency (Figure 4). For monocular within-scale masking, the additional mask frequencies increase masking but only when phase aligned. For monocular viewing, cross-scale masking that is not dependent on phase occurs. For dichoptic viewing, the additional mask components act to reduce the masking and the extent of this reduction depends on phase. In summary, our results show that not only do frequencies far away from the target have an influence but also this influence depends on their phase structure.

There seems to be at least two mechanisms of interocular suppression. One is suppression triggered by the phase alignment. Maehara et al. (2009) argued that phase alignments signal the existence of a luminance contour in the monocular image and that this signal suppresses information conveyed by the other eye when there is no corresponding contour signal. On the other hand, according to Meese, Challinor, and Summers (2008), monocular and dichoptic cross-orientation suppression involve suppression pathways that are not active under binocular viewing. Since their stimuli were composed of one or two gratings, it would be reasonable to assume that there is another type of interocular suppression that is effective without phase alignment. Meese et al. (2008) explained this suppression in terms of changes in linearity of contrast gain control. There were stimulus differences in orientations, spatial frequencies, and phases between their experiments and ours. It remains to be examined which stimulus properties are critical for each type of interocular suppression.

Are these results explicable in terms of Foley’s (1994) model of pattern masking?

In Foley’s (1994) model, inhibitory components are generated independently and linearly from the different stimulus component contrasts within an orientation channel: The inhibitory signal within an orientation channel is a linear function of the component contrasts, raised to the power of q. Since our Q (or Qp) mask is the sum of S and M (or Mp) masks, Foley’s model predicts that the inhibitory signal due to the Q and Qp masks should be predictable from those of S and M, and Mp. We tested this prediction using the modeling procedure described in Appendix A; this was done for the monocular viewing conditions only, because there is no agreed model at present for binocular combination that can accommodate different stimulus phases and spatial frequencies. Figure A2 shows performance of the model fitted to the S, M, and Mp conditions simultaneously and the predictions for the Q and Qp conditions using the fitted parameters (parameter values and model performance are given in Table A1). Figures A3 and A4 show performance of the model fitted separately to the phase-aligned (S and M) and phase-scrambled (S and Mp) conditions, respectively (parameter values and model performance are given in Table A2). The model fitted to S, M, and Mp (Figure A2) or just S and Mp (Figure A4) is able to predict the Qp data fairly well, but the model does not provide a satisfactory prediction of the Q data, even when the model is fitted just to the S and M conditions (Figure A3): The thresholds are higher than predicted from the independence assumption of Foley’s model. The additional masking that occurs for the square wave shows that phase alignment between the stimulus components can increase the inhibitory signal for monocular masking, so that the masking effect is greater than we would expect if the inhibitory signals from different components of a stimulus were generated independently.

Is dichoptic blur suppression explicable in terms of low-level phase-dependent masking?

An obvious candidate for dichoptic blur suppression would be low-level pattern masking in which phase-aligned higher harmonics (which produce a sharp edge representative of a sharply focused image) presented to one eye would mask a lower spatial frequency sinusoid (with blurry edges representative of a blurred image) seen by the other. We measured the dichoptic masking effect of phase-aligned higher harmonics (Q and M masks) on sine-wave targets. We did not find the expected stronger masking compared with a sine-wave mask under dichoptic viewing. The dichoptic masking produced by a square-wave mask (i.e., sharp in focus) on a sine-wave target is equal to or less than that of a sine-wave mask (i.e., blurry stimulus) on a sine-wave target. This cannot provide a low-level masking explanation for why a sharp stimulus in one eye is so effective at suppressing visibility of a blurred stimulus seen in the other (i.e., so-called dichoptic blur suppression).

The broadband nature of pattern masking

Previous work points to the broadband nature of masking. Using a suprathreshold discrimination paradigm, Sagi and Hochstein (1984) concluded that at high contrast, much broader spatial frequency mechanisms are involved than at threshold, with bandwidths ranging between 1.6 and 3.2 octaves. In a noise masking study of motion and stereo sensitivity, Hess, Wang, and Liu (2006) argued that for some tasks information is not accessible from narrowband spatial frequency mechanisms and that the resultant tuning can exceed 3 octaves. Recently, Henning and Wichmann (2007) suggested that the pedestal effect in pattern masking may stem from the use of contrast information carried by spatial mechanisms tuned to very different spatial frequencies from that of the target pattern. Their follow-up study proposed a multichannel model for pattern masking (Goris et al., 2009).
The phase dependence of pattern masking and its relationship to natural scenes

There is supporting evidence pointing to the importance of phase alignments in masking. A case has been made for the importance of phase alignments for the discrimination of faces (Hansen, Farivar, Thompson, & Hess, 2008) and natural scenes (Hansen & Hess, 2007) and for their dominance in masking (Bex et al., 2009). Bex et al. (2007) found that the TVcC functions in natural scenes changed with phase structure at remote spatial scales. They also made a similar comparison to that in the current study between sine-wave, square-wave, and missing fundamental stimuli (Bex et al., 2009). Their comparison was as a function of spatial frequency (fixed contrast), whereas ours was as a function of contrast (fixed frequency). Their results, particularly observer CV (their Fig. 10), show significant but reduced masking for the fundamental of the square wave (comparable to that for the missing fundamental waveform). They attribute this masking at the “missing fundamental frequency” of the waveform to the edge structure. In terms of dichoptic masking, it has previously been argued (Maehara et al., 2009) that, for high-pass maskers, phase alignments are crucial. Their approach was different to ours in that they used an additional mask as well as a target/pedestal paradigm. To the extent to which we can compare our results with theirs, we also show that, for masks composed of higher spatial frequencies than that of the target, dichoptic masking is much greater for phase-aligned components.

The phase dependence of dichoptic masking and its relationship to binocular rivalry

Dichoptic masking and binocular rivalry have been used to examine interocular interactions. It has been shown that sharp images dominate in rivalry over blurred images (either done with phase scrambling or low-pass filtering; Arnold, Grove, & Wallis, 2007; Baker & Graf, 2009; Fahle, 1982). Although we did not observe any rivalry under our stimulus presentations, our Q mask did not produce stronger dichoptic masking than the S mask, suggesting that dichoptic masking and binocular rivalry may involve different types of interocular interaction (Legge, 1979).

Conclusion

We used a pattern masking paradigm to assess the effects of mask spatial frequencies that were at least 3 times the target frequency and their phase alignment. Our results showed different patterns of performance between monocural and dichoptic viewing conditions. Under monocular viewing conditions, phase alignment makes a big difference to the effect of the Q vs. Qp mask but not M vs. Mp mask. These results suggest that masking mechanisms are more complicated than previously thought. To explain all of our results, we would need the following: pure cross-scale masking (M and Mp masks) that is phase insensitive but also remote spatial frequency channels that can modulate the phase-sensitive, within-scale masking (Q and Qp masks). Under dichoptic viewing conditions, strong dichoptic masking occurs when information is matched in scale between the eyes, and phase alignment plays a more crucial role. Surprisingly, adding phase-aligned harmonics to a dichoptic S mask (thereby sharpening the edges in the mask and increasing its contrast energy) diminishes its masking effect.

Appendix A

Divisive inhibition model

We attempted to account for our results using Foley’s (1994) divisive inhibition model, which assumes that, within an orientation channel, the inhibitory signal is a linear function of the contrasts of the different spatial frequency components, raised to the power of $q$. The model assumes that there are filters matched to the target component and to the other stimulus components. Because the first stage of Foley’s model involves linear summation of stimulus components within an orientation channel (see Foley’s Eq. 7), we can conceptually collapse the set of filters down to just two: one narrowband filter that is sensitive to the target (the “detection filter”) and another that is sensitive to all the other stimulus components; the latter would be constructed from the filters that respond to all the non-target components but can be considered to be a single, broadband, filter. We assume that detection performance is mediated by the narrowband “detection filter.” It might be considered more appropriate to use Foley and Chen’s (1999) extension of Foley’s model because only Foley and Chen’s model includes filters selective for phase. However, because our mask components are always either in phase with the target or outside of the spatial frequency passband of the detection filter (or both), Foley and Chen’s model is formally equivalent to Foley’s model for our phase-aligned conditions and almost equivalent to it for the phase-randomized conditions. We assume that the detection filter is only sensitive to the S stimulus and the S component of Q. The excitatory responses, $E_s$, $E_m$, and $E_q$, of the detection filter in the S, M, and Q conditions, respectively, are given by

$$E_s = S e_s (c + \Delta c),$$  \hspace{1cm} (A1)  

$$E_m = S e_s \Delta c,$$  \hspace{1cm} (A2)  

$$E_q = S e_s (c + \Delta c),$$  \hspace{1cm} (A3)

where $S e_s$ is the detection filter’s excitatory sensitivity to the S component, $c$ is the mask contrast, and $\Delta c$ is zero for
Figure A1. Schematic illustration of the divisive inhibition model.

Figure A2. Monocular viewing results with model fitted to S, M, and Mp conditions simultaneously (blue and green) and predictions for the Q and Qp conditions (red and pink). The leftmost columns show the phase-aligned results, while the rightmost columns show the phase-scrambled results. Green filled circles and lines denote the data and model fits for a sine-wave grating mask, and blue filled diamonds and lines denote the data and model fits for a missing fundamental mask. The second column shows the results for a square-wave mask (red squares) and the red line denotes the prediction from the divisive inhibition model that assumed that the inhibitory signal for the Q stimuli is equal to the linear sum of those for S and M. The extreme right column shows the corresponding results for the phase-scrambled configuration and the layout is the same as phase-aligned configuration.
The inhibition input went through a similar process so the inhibitory inputs, $I_S$ and $I_M$, for the S and M conditions, respectively, are given by

$$I_S = S_I S (c + \Delta c), \quad (A4)$$

$$I_M = S_{IM}, \quad (A5)$$

where $S_{IS}$ and $S_{IM}$ are the inhibitory sensitivities of the narrowband detection filter and the broadband filter, respectively. Because of the linear summation of inhibitory components across spatial frequency within an orientation channel in Foley’s (1994) model, the inhibitory input, $I_Q$, for the Q mask (which is the sum of S and M masks) should be the linear sum of inhibitory inputs for the other two conditions:

$$I_Q = I_S + I_M. \quad (A6)$$

The response of the detection filter is then given by

$$R = \frac{E^p}{\mu + \sigma}, \quad (A7)$$

where $\sigma$ is an additive constant, and $E$ is $E_S$, $E_M$, or $E_Q$ depending on the masking condition and similarly $R$ is $R_S$, $R_M$, or $R_Q$. Suppose the threshold is determined by the difference between the response to the mask alone ($R_m$, when $\Delta c = 0$) and the mask plus target ($R_{m+t}$, when $\Delta c$ is the target contrast). The threshold, $\Delta c$, is reached when the expected difference in response

$$D = R_{m+t} - R_m, \quad (A8)$$

reaches unity (Figure A1).

First, we fitted data from the S, M, and Mp conditions simultaneously, using different parameters $S_{IM}$ and $S_{IM}$ for the inhibitory sensitivities in the M and Mp conditions, respectively. Then, the fitted parameters were used to predict the performance for Q and Qp masks based on the assumption that the broadband filter response to the Q (or Qp) mask is a linear function of the responses to the S and M (or Mp) masks (Equation A6). We constrained our parameters—$p$, $q$, and $\sigma$—to be the same for all conditions. Thus, the model has 7 parameters ($S_{ES}$, $S_I S$, $S_{IM}$,

<table>
<thead>
<tr>
<th>PCH</th>
<th>TD</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_{E_S}$</td>
<td>100</td>
</tr>
<tr>
<td>$S_{IS}$</td>
<td>57.03</td>
</tr>
<tr>
<td>$S_{IM}$</td>
<td>9.02</td>
</tr>
<tr>
<td>$S_{IM}$</td>
<td>9.79</td>
</tr>
<tr>
<td>$p$</td>
<td>2.36</td>
</tr>
<tr>
<td>$q$</td>
<td>1.89</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>7.89</td>
</tr>
<tr>
<td>Numobs</td>
<td>19</td>
</tr>
<tr>
<td>$\chi^2_{\text{red}}$</td>
<td>6.919 ($p = 0.863$)</td>
</tr>
<tr>
<td>$x^2_Q$</td>
<td>24.402 ($p &lt; 0.01$)</td>
</tr>
<tr>
<td>$x^2_{Qp}$</td>
<td>9.123 ($p = 0.167$)</td>
</tr>
</tbody>
</table>

Table A1. The parameters of model fitted to S, M, and Mp conditions simultaneously and tests of statistical significance of the model predictions. Notes: *Fixed parameter. Numobs: Number of data points used to fit the data. $\chi^2_{\text{red}}$: The reduced chi-squared statistics were calculated according to the following formula,

$\chi^2_{\text{red}} = \frac{1}{n} \sum \frac{(\text{obs}_i - \text{pred}_i)^2}{\text{var}_i}$,

where $n$ is the number of degrees of freedom, given by Numobs $- n - 1$, with $n$ being the number of fitted parameters, and $\sigma$ is the known variance of the observation (here, we used averaged variance). $x^2_{Q,Mp}$: The same equation as previously mentioned. The degrees of freedom are 7–1.

non-target stimuli and equal to the target contrast for target stimuli.

The inhibition input went through a similar process so the inhibitory inputs, $I_S$ and $I_M$, for the S and M conditions, respectively, are given by

$$I_S = S_I S (c + \Delta c), \quad (A4)$$

$$I_M = S_{IM}, \quad (A5)$$

where $S_{IS}$ and $S_{IM}$ are the inhibitory sensitivities of the narrowband detection filter and the broadband filter,
Si\textsubscript{Mp}, \( p \), \( q \), and \( \sigma \), with \( S_e \) being fixed at 100 and 6 parameters free to vary. The fitted results are shown in Figure A2 and the parameters and the \( x^2 \) goodness-of-fit test are shown in Table A1. The divisive inhibition model fits the S, M, and Mp mask results quite well. Although the estimated parameters predict the performance of Qp mask fairly well, the derived parameters underestimated the masking effect found with the Q mask.

We then fitted data from the phase-aligned and phase-scrambled configuration separately and again tried to predict the performance of Q or Qp masks according to the phase-aligned or phase-scrambled parameters, respectively. The model has 6 parameters for each fit (\( S_e \) was fixed at 100, and 5 parameters were free to vary—see Table A2). The fitted results for the phase-aligned configuration are shown in Figure A3 and the parameters and the \( x^2 \) goodness-of-fit test are shown in Table A2. The divisive inhibition model fit the S and M mask data quite well. However, the derived parameters underestimated the masking effect found with the Q mask even though the model was fitted just to the phase-aligned data. The fitted results for the phase-scrambled configuration are shown in Figure A4 and the parameters and the \( x^2 \) goodness-of-fit test are shown in Table A2. The divisive inhibition model fit the S/Mp mask data quite well and the estimated parameters predict the performance of Qp mask better than in the phase-aligned case. We also tested a non-linear summation model, where the inhibitory denominator involves non-linear summation, namely, the output of the \( I_S \) and \( I_M \) were summed together after each of them were raised by power \( q \). However, the goodness of fit was worse (\( x^2 \) are 26.788, 52.834, and 22.904 for PCH, TD, and GM, respectively, for phase-aligned configuration; \( x^2 \) are 15.194 and 2.233 for PCH and TD, respectively, for phase-scrambled configuration). The modeling results suggest that phase alignment of mask components must play an important role in pattern masking.

**Acknowledgments**

This research is funded by the NSERC Grant (RGPIN 46528-06) to RFH. KAM is supported by EPSRC Grant EP/H033955/1 to Joshua Solomon. GM is supported by the Grant in Aid for JSPS Fellows from the Japan Society for Promotion of Science.
References


