Spatial and temporal dependencies of cross-orientation suppression in human vision

Tim S. Meese* and David J. Holmes

School of Life and Health Sciences, Aston University, Birmingham B47ET, UK

A well-known property of orientation-tuned neurons in the visual cortex is that they are suppressed by the superposition of an orthogonal mask. This phenomenon has been explained in terms of physiological constraints (synaptic depression), engineering solutions for components with poor dynamic range (contrast normalization) and fundamental coding strategies for natural images (redundancy reduction). A common but often tacit assumption is that the suppressive process is equally potent at different spatial and temporal scales of analysis. To determine whether it is so, we measured psychophysical cross-orientation masking (XOM) functions for flickering horizontal Gabor stimuli over wide ranges of spatio-temporal frequency and contrast. We found that orthogonal masks raised contrast detection thresholds substantially at low spatial frequencies and high temporal frequencies (high speeds), and that small and unexpected levels of facilitation were evident elsewhere. The data were well fit by a functional model of contrast gain control, where (i) the weight of suppression increased with the ratio of temporal to spatial frequency and (ii) the weight of facilitatory modulation was the same for all conditions, but outcompeted by suppression at higher contrasts. These results (i) provide new constraints for models of primary visual cortex, (ii) associate XOM and facilitation with the transient magno- and sustained parvocortical, respectively, and (iii) reconcile earlier conflicting psychophysical reports on XOM.

Keywords: contrast gain control; psychophysics; cross-orientation inhibition; masking; human vision

1. INTRODUCTION

Cross-orientation suppression (XOS) is the process whereby a cell's response to a stimulus at its preferred orientation is reduced by the superposition of a mask stimulus at another orientation (Bonds 1989; Nelson 1991; Allison *et al.* 2001), often perpendicular to the preferred stimulus (Bishop *et al.* 1973; Morrone *et al.* 1982; DeAngelis *et al.* 1992). This has attracted considerable attention owing to its supposed involvement in orientation tuning (Sillito 1979; Bonds 1989; Somers *et al.* 1995; Seriès *et al.* 2004), contrast gain control (Albrecht & Geisler 1991; DeAngelis *et al.* 1992; Heeger 1992; Carandini *et al.* 1997), perceptual grouping and segmentation (Morrone *et al.* 1983), adaptation after-effects (Snowden & Hammett 1992; Barrett *et al.* 2002) and efficient image coding strategies (Bonds 1991; Schwartz & Simoncelli 2001; Field & Wu 2004; Olshausen & Field 2005; though see Seriès *et al.* 2004).

A psychophysical phenomenon that has been closely linked to XOS is cross-orientation masking (XOM). However, clear evidence for XOM was slow to emerge. Vision research in the 1970s and 1980s was dominated by the notion of independent spatial channels (Campbell & Robson 1968; Blakemore & Campbell 1969; Graham 1989), and masking was supposed to occur only when the mask excited the mechanism used to detect the target stimulus (Wilson *et al.* 1983; Daugman 1984; Phillips & Wilson 1984; Anderson & Burr 1989; Harvey & Doan 1990). In this model, an orthogonal mask would not influence the orientation-tuned filters of visual cortex, and this is what was reported by several early studies

(Campbell & Kulikowski 1966; Daugman 1984; Harvey & Doan 1990). However, a subsequent erosion of the independent channels model (e.g. Henning *et al.* 1975; Olzak 1985; Burr & Morrone 1987; Olzak & Thomas 1991) was followed by an influential proposal involving mutual inhibition between filters tuned to the full range of orientations (Albrecht & Geisler 1991; Heeger 1992). This model of cortical physiology prompted more detailed psychophysical measurements of contrast-masking functions (Foley 1994), which revealed that a cross-oriented mask could raise detection threshold (quite substantially) after all (Ross & Speed 1991; Ross *et al.* 1993; Foley 1994). This process is now commonplace in functional models (Foley & Chen 1997; Thomas & Olzak 1997; Graham & Sutter 1998; Olzak & Thomas 1999; Meese 2004; Baker *et al.* submitted) and image-processing models of suprathreshold spatial vision (Watson & Solomon 1997; Brady & Field 2000; Itti *et al.* 2000; Clatworthy *et al.* 2003). Nevertheless, the question remains why earlier psychophysical masking studies failed to find XOM. Furthermore, authors of most contemporary models of the process make a simplifying, often tacit, assumption that XOS is scale invariant (Graham & Sutter 1996; Rohaly *et al.* 1997; Watson & Solomon 1997; Brady & Field 2000; Clatworthy *et al.* 2003), but there is no evidence to support this.

We address these issues here by measuring contrast-masking functions of XOM¹ for a wide range of spatio-temporal frequencies. We find that XOM is a low spatial frequency (SF) and high temporal frequency (TF) phenomenon. This places new constraints on models of spatial vision as well as reconciling the literature on XOM in humans (see §4).

* Author for correspondence (t.s.meese@aston.ac.uk).

2. MATERIAL AND METHODS

(a) *Equipment and observers*

Stimuli were displayed on a Sony monitor using a VSG2/3 or VSG2/4 stimulus generator controlled by a PC. Mean luminance of the central display region was 62 cd m^{-2} . The display was gamma corrected to ensure linearity over the full contrast range, and a frame interleaving technique (60 Hz per image) allowed the contrasts of the mask and the test stimuli to be varied independently using lookup tables. The VSG2/3 and VSG2/4 were operated in pseudo 12-bit mode and pseudo 15-bit mode, respectively.

The observers were the two authors (D.J.H. and T.S.M., aged 30 and 40, respectively). Both are psychophysically experienced, wore their normal optical correction throughout and viewed the stimuli binocularly. They made observations in a dark room with their heads in a chin and forehead rest.

(b) *Stimuli*

Mask and test stimuli were sine-phase, circular Gabor patches. The test patch was always horizontal and the mask patch was always vertical. With the exception of contrast, the mask and the test were otherwise identical (figure 1). The Gabor patches were scaled with SF such that their full-width at half-height was always 1.65 cycles. This involved manipulating both the viewing distance and the stimulus dimensions rendered on the display. Carrier spatial frequencies were 0.5, 1, 2, 4 and 8 c deg^{-1} , and were viewed from 57, 114, 114, 228 and 228 cm, respectively. Four different temporal waveforms were used (figure 2). These had peak temporal frequencies of 0.5 (nominal), 1, 4 and 15 Hz (see the caption of figure 2 and §4 for further details and comment). Stimulus contrast is expressed as Michelson contrast in per cent, given by $c = 100 (L_{\max} - L_{\min}) / (L_{\max} + L_{\min})$, and in dB re 1% ($20 \log(c)$). Mask contrast was in the range of 0–45% (33 dB).

In detection experiments, it is common practice to place a small fixation point in the centre of the display. However, we were concerned that this might have had a suppressive influence on our stimuli, particularly at the higher spatial frequencies. To try and avoid this problem, we used a ‘quad’ of points arranged around the stimulus patch in a square. These points were always three carrier cycles away from the centre of the display and were visible throughout the experimental sessions. This arrangement provided a very strong cue to stimulus location with the aim of reducing stimulus uncertainty (Pelli 1985; Petrov *et al.* 2006).

Later we shall describe our stimuli in terms of their speed (the ratio of temporal and spatial frequencies). Although our stimuli did not drift, a flickering grating is equivalent to the superposition of a pair of identical grating components drifting in opposite directions. The spatio-temporal frequency ratio is the speed of these components.

(c) *Procedure*

A two-interval forced-choice (2IFC) procedure was used to estimate detection thresholds. One interval contained only the mask patch and the other contained the mask plus the test patch. Observers indicated which interval contained the test patch by pressing one of the two mouse buttons and auditory feedback indicated correctness of response. Stimulus conditions were ordered by random selection from nested blocks of TF, mask contrast and SF. For each condition, a pair of randomly interleaved three-down, one-up staircases was used with a step size of 2.5 dB. Each staircase terminated after 12 reversals and the data from the pair collapsed before

Figure 1. Mask and test Gabor patches used in the experiment. For D.J.H., the carrier spatial frequencies were 0.5, 1, 2, 4 and 8 c deg^{-1} . For T.S.M., the spatial frequencies were 1 and 4 c deg^{-1} . All the stimuli had a full width at half-height of 1.65 cycles.

performing probit analysis to estimate the threshold (75% correct) and the standard error. In the rare cases where the standard error was greater than 3 dB, the data were discarded and the session was rerun. All the conditions were performed three times, resulting in estimates from about 300 experimental trials per threshold.

The time-intervals between (i) 2IFC stimulus presentations and (ii) the release of the button press and the onset of the next stimulus were designed to avoid interference from afterimages. These differed across conditions and observers and are summarized in table 1.

3. RESULTS

Figure 3 shows detection threshold as a function of mask contrast for 15 different spatio-temporal frequencies for D.J.H. XOM is the most severe in the bottom left corner of the figure, where TF is high and SF is low. Towards the opposite corner of the figure (high SF and low TF), there is very little evidence for XOM. In fact, in this region (and elsewhere) there are several instances where the mask causes the threshold to drop below that measured at baseline (mask contrast = 0%; horizontal dashed lines), indicating cross-orientation facilitation (XOF) of the order of 2 or 3 dB. We shall consider this unexpected phenomenon further in §4.

One possibility is that XOM is not seen at the higher spatial frequencies because the loss of sensitivity in that region means that the mask does not rise to sufficient potency. To show that this was not so, the entire dataset was replotted in figure 4, where the mask and test contrasts have been normalized to detection threshold for each spatio-temporal frequency (we assume that sensitivity to mask and test is the same). The level of masking differs across the three panels (different temporal frequencies) and also within each panel (different spatial frequencies), providing good evidence that XOM depends on both TF and SF. Therefore, XOM is not scale invariant in either space or time. This can be illustrated with a specific example in the SF domain. As an observer moves away from a pair of flickering orthogonal gratings, the retinal SF of the image will increase, but the suppressive influence between the gratings will decrease.

(a) *A functional model of XOM and XOF*

Here, we develop a widely used functional expression for contrast gain control (Foley 1994; Meese 2004) to accommodate XOF and to fit our masking data.

Figure 2. Temporal waveforms used in the experiment. (a) The nominal 0.5 Hz waveform is a raised negative cosine function with a central plateau of 500 ms. (b) The 1 Hz waveform is a single cycle of a sine-wave. (c) The 4 Hz waveform is two cycles of a sine-wave multiplied by a raised negative cosine function with a central plateau of 250 ms (one cycle). (d) The 15 Hz waveform is a biphasic pulse with a length of four frames (33.3 ms = two frames each for mask and test) in each phase. For D.J.H., the temporal frequencies were 0.5, 4 and 15 Hz. For T.S.M., the temporal frequencies were 1, 4 and 15 Hz.

Table 1. Timing parameters for stimulus presentations. ('2IFC ISI' is the period between the two stimulus intervals in a trial. 'Post-response' is the duration between the offset of a button press and the onset of the next stimulus presentation.)

temporal frequency (Hz)	stimulus duration (ms)	T.S.M.		D.J.H.	
		2IFC ISI (ms)	post-response (ms)	2IFC ISI (ms)	post-response (ms)
0.5	1000	not tested	not tested	1000	1000
1	1000	700	300	not tested	not tested
4	500	400	200	1000	1000
15	67	400	200	1000	1000

Figure 3. XOM functions for D.J.H. Horizontal lines indicate detection thresholds for a mask contrast of 0%. Data points above and below these lines indicate masking and facilitation, respectively. Different panels are for different spatio-temporal frequencies (see figure labels). Error bars show ± 1 s.e.m. The data are from approximately 31 500 trials.

Figure 4. Normalized XOM functions for D.J.H. Mask contrast of 0 dB corresponds to detection threshold for the test patch on both axes. Mask contrasts below 0 dB are sub-threshold. TF is different for the three panels (insets), and different symbols are used for different SFs (see legends). Error bars show ± 1 s.e.m.

The response (resp) of the detecting mechanism is given by

$$\text{resp} = \frac{c_{\text{test}}^p (1 + \alpha c_{\text{mask}})}{1 + c_{\text{test}}^q + w c_{\text{mask}}^q}, \quad (3.1)$$

where c_{test} and c_{mask} are test and mask contrasts (in %) and p , q , w and α are free parameters. The exponents p and q are usually constrained by pedestal masking experiments, but no such experiments were performed here. Instead, we set them to widely used defaults of $p=2.4$ and $q=2.0$ (Legge & Foley 1980). Their precise values are not important for the quality of our fits.

In 2IFC, the decision variable at threshold is given by the difference between the detecting mechanism's response to the mask alone and mask plus test

$$\text{resp}_{\text{test+mask}} - \text{resp}_{\text{mask}} = k, \quad (3.2)$$

where k is proportional to the standard deviation of late additive noise in the model. However, the model is to be fit to data that are normalized to detection thresholds which, from equations (3.1) and (3.2), constrains $k=0.5$ ($(c_{\text{test}}/c_{\text{test}})^p / (1 + c_{\text{test}}/c_{\text{test}})^q - 0 = k = 1/2$). This leaves two free parameters (w and α) for each masking function. The weight, w , controls the suppressive influence of the mask on the denominator which is sometimes called XOS and is designed to produce XOM. In equation (3.1), the numerator term in parenthesis provides a modulatory influence from the mask on excitation, producing cross-orientation enhancement (XOE). This is designed to accommodate XOF and is controlled by the free parameter α . Similar arrangements, but for different contexts, have been considered by Xing & Heeger (2001), Chen & Tyler (2002) and Yu *et al.* (2003). Note that XOS is subject to the accelerating exponent q , whereas XOE is not. This means that suppression will outpace enhancement as contrast increases.

We fit the model to the full dataset from D.J.H. using numerical analysis to solve the model equations for c_{test} and a downhill simplex algorithm to find the best fitting parameters by minimizing the RMS error between the data and the model fit. The RMS error for the entire dataset (15 functions, 90 data points and 30 free parameters) was 0.62 dB (not shown) and is very good. To try and understand the relation between the free parameters and the spatio-temporal characteristics of the

stimuli, we plotted α and w as functions of SF, TF and speed (where speed = TF/SF). We found no systematic relation for α and matters were little better for w , though a power function with exponent of 0.31 accounted for 19% of the variance in a plot of w against speed. Next we reduced the number of free parameters in the model by constraining w or α to have the same value across the entire dataset. Fixing w in this way was a dismal failure, but fixing α caused the RMS error of the fit to decline by only 0.19 dB and saved 14 free parameters. The fits are shown in figure 5a for D.J.H., where 16 free parameters describe 15 masking functions. A plot of w against speed (TF/SF) now reveals lawful behaviour with a power law accounting for 89% of the variance (figure 6a). We performed the same analysis on the results of T.S.M., which had a similar form to those of D.J.H. For this observer, six masking functions were fit by seven free parameters with an RMS error of 0.77 dB (figure 5b), and a power law accounted for 98% of the variance of w against speed (figure 6b). For T.S.M., the weight of XOF (α) was less than that of D.J.H. (insets in figure 6).

4. DISCUSSION

(a) Physiological constraints

We performed XOM experiments and found that the weight of XOS (w) depends upon both SF and TF in a lawful way. To a first approximation, the relation is $w \propto \sqrt{\text{TF}/\text{SF}}$. Why the dependency should have this exact form is not clear, but we suggest that it might be related to constraints of the underlying physiology. It is generally agreed that the magnocellular stream extends the visual system's temporal resolution and that the parvocellular stream extends its spatial resolution (Merigan & Eskin 1986; Merigan *et al.* 1991; Spear *et al.* 1994; Lennie & Movshon 2005). Thus, M- and P-cells are the most active at fast and slow stimulus speeds, respectively. It is also well known that in the retina and lateral geniculate nucleus, P-cells are fairly linear, whereas M-cells saturate (Derrington & Lennie 1984; Kaplan & Shapley 1986; Solomon & Lennie 2005), and there is some evidence that this extends to at least some striate cells (Sclar *et al.* 1990; Solomon & Lennie 2005). One consequence of saturation is that it interferes with the pattern of activity across arrays of cells that sample various image dimensions. (These patterns are sometimes known

Figure 5. Model fits and normalized XOM functions of (a) D.J.H. and (b) T.S.M. Note the different SFs and TFs used by the two observers. In each panel, 0 dB corresponds to detection threshold for a mask contrast of 0% (baseline), as shown by the left most point in each plot. The model was constrained to pass through these points, and so they did not contribute to the calculation of RMS errors of the fits. These were 0.81 and 0.77 dB for D.J.H. and T.S.M., respectively. Error bars show ± 1 s.e.m.

as population codes.) The problem is that strong signals drive much of the population into saturation and this distorts the signal's signature. However, by using the overall population activity to normalize the response of

each cell in the population, saturation occurs at lower signal strengths for the less strongly driven cells and this preserves the signal's signature (Albrecht & Geisler 1991; Heeger 1992). Thus, the saturating magnocellular

Figure 6. Relation between the free model parameter w (the weight of XOS) and stimulus speed (TF/SF) of (a) D.J.H. and (b) T.S.M., with α set as shown in the insets. Such lawful behaviour was not found when w was plotted as a function of SF or TF.

pathway has a much stronger need for the normalizing effects of XOS than does the more linear parvocellular pathway. This is consistent with our finding that XOS is a high-speed (magnocellular) phenomenon.

Another issue relating to anatomy and physiology concerns the locus of XOS. Heeger (1992) and others envisaged a single mechanism of intracortical inhibition, but recent single-cell work in cat suggests that at least two pathways are involved. One of these is cortical and dichoptic (Li *et al.* 2005; Sengpiel & Vorobyov 2005), and the other is sub-cortical and monoptic (Freeman *et al.* 2002; Bonin *et al.* 2005; Li *et al.* 2005; Sengpiel & Vorobyov 2005). We have also found evidence for at least two suppressive pathways in human, consistent with the anatomical arrangement proposed for cat (Baker & Meese 2006; Baker *et al.* submitted). For simplicity, the model developed for our binocular stimuli here has no specific routes for eye of origin and only a single (binocular) route for XOS. This arrangement has been very successful in binocular-masking studies (Legge & Foley 1980), including those that have addressed binocular XOM (Ross & Speed 1991; Foley 1994; Watson & Solomon 1997; Meese 2004). Nevertheless, we envisage that the current model will need to be further developed to handle the details of cross-oriented monoptic and dichoptic masking in the spatio-temporal frequency domain.

(b) Temporal waveforms

Of the four temporal waveforms that we used (figure 2), three waveforms at 1 Hz and above are band-limited. However, the nominal 0.5 Hz stimulus (used only by D.J.H.) has no complementary negative lobe and its amplitude spectrum peaks at 0 Hz. Of course, most of the stimulus energy is above 0 Hz, and so we chose a nominal label of 0.5 Hz, on the grounds that the positive lobe represents half of one cycle (1000 ms). But by expressing our stimuli in terms of their speed (TF/SF), this decision is not arbitrary. Fortunately, we found that our conclusions are robust against the specific value chosen. We recalculated the slopes of the power functions in figure 6a, reassigning the 0.5 Hz conditions first to 0.25 Hz and then 1 Hz, and found log-log slopes (exponents) of 0.41 and 0.55, and r^2 values of 0.85 and 0.92, respectively. The deviations between these figures and the values associated with figure 6a (slope = 0.48; r^2 = 0.89) are marginal and do not change our conclusions.

(c) Meta-analysis

As mentioned in §1, early psychophysical masking studies found no threshold elevation from a cross-oriented mask, whereas later studies did. In our present study, we have found that XOM depends upon both SF and TF, and these parameters have varied widely between previous studies. Table 2 summarizes this variation as well as our estimates of the stimulus speeds (TF/SF) used and whether XOM was found. It is clear that speed accounts for much of the discrepancy in the earlier literature (see table 2 caption for further details of the analysis).

A previous study of note is that of Burbeck & Kelly (1981), who measured detection thresholds (using a method of adjustment) in the presence of cross-oriented masks for a wide range of spatio-temporal frequencies. Like us, they found that XOM was located in the high-speed corner of spatio-temporal space. However, those authors interpreted their results in terms of a pedestal masking process within non-oriented filters. Ferrera & Wilson (1985) made the same interpretation of their own data, but the idea is not consistent with results from twin-masking studies (Ross & Speed 1991; Foley 1994; Holmes & Meese 2004) or analyses of the slope of the psychometric function (Georgeson & Meese 2004). For example, if the mask stimulus excites the detecting mechanism (as for pedestal masking), then for a monotonic transducer, the input/output characteristic is approximately linear for small changes in input. Consequently, the psychometric function will have a d' exponent of unity, equivalent to a Weibull $\beta \sim 1.3$ (Pelli 1985, 1987; Bird *et al.* 2002). We estimated the slopes of the psychometric functions measured here for D.J.H. (the geometric mean of β across three sessions; approximately 300 trials per function) and found that none were as shallow as $\beta = 1.3$. In fact, for only 12 out of 105 psychometric functions, $\beta < 2$ and the median $\beta = 2.95$. It is very unlikely that our XOM results are caused by within-channel excitation by the mask.

The spatio-temporal properties of XOM have implications for studies that have used contrast masking to measure the bandwidth of detecting mechanisms. Phillips & Wilson (1984) estimated that the orientation bandwidth of low SF mechanisms was considerably greater than those at higher SFs, but their model took no account of XOS. In another work (Meese & Holmes 2003), we have reexamined those data and similar

Table 2. Meta-analysis of psychophysical XOM studies for non-drifting stimuli. (The table is ordered according to whether XOM was found and then chronologically. The study by Burbeck & Kelly (1981) appears three times, owing to the wide range of conditions they investigated: see their work for further details and further conditions. In studies where fixed stimulus duration and no flicker were used (e.g. duration = 500 ms), TF was thus estimated as $TF \sim 1/(2 \times \text{duration})$. For convenience, we label temporal frequencies less than or equal to 0.5 Hz as 'slow', those greater than or equal to 2 Hz as 'fast' and those in between as 'intermediate'. Whether XOM was found is clearly dependent upon stimulus speed (MOA is method of adjustment, and *n*IFC is '*n*' interval forced-choice).)

evidence for XOM	speed (TF/SF, deg s ⁻¹)	SF (c deg ⁻¹)	TF or duration	study	method
no	slow (<0.05)	10	long	Campbell & Kulikowski (1966)	MOA
no	slow (≤ 0.33)	≥ 3	≤ 1 Hz	Burbeck & Kelly (1981)	MOA
no	slow (~ 0.125)	8	500 ms	Daugman (1984)	2IFC
no	slow (~ 0.125)	8	500 ms	Harvey & Doan (1990)	3IFC
intermediate	intermediate (~ 1.3)	~ 1.5	~ 2 Hz	Burbeck & Kelly (1981)	MOA
intermediate	slow (~ 0.5)	4	250 ms	Itti <i>et al.</i> (2000)	2IFC
yes	fast (≥ 4)	≤ 1	≥ 4 Hz	Burbeck & Kelly (1981)	MOA
yes	fast (≥ 4)	≤ 2	8 Hz	Ferrera & Wilson (1985)	1IFC
yes	fast (4.4)	2	8.8 Hz	Ross & Speed (1991)	2IFC
yes	fast (4.4)	2	8.8 Hz	Ross <i>et al.</i> (1993)	2IFC
yes	fast (2)	1	2 Hz	Snowden (1994)	2IFC
yes	fast (~ 7.6)	2	33 ms	Foley (1994)	2IFC
yes	intermediate (~ 0.83)	2	300 ms	Foley & Chen (1997)	2IFC
yes	fast (7.5)	2	15 Hz	Meese & Holmes (2002)	2IFC
yes	fast (~ 5)	1	100 ms	Meese (2004)	2IFC
yes	intermediate (~ 0.67)	5	150 ms	Petrov <i>et al.</i> (2005)	2IFC

(unpublished) data of our own, with a model incorporating oriented filters and XOS. At 15 Hz and 0.5 c deg^{-1} , our preliminary analysis suggests an orientation bandwidth that is narrower than that often supposed (approx. $\pm 17^\circ$ at half height) and that decreases only gently over a three-octave range. With constant TF (15 Hz), this increase in SF results in a decrease in speed. We suggest that in other studies, much of the masking attributed to within-channel influences from the mask on broadband mechanisms was in fact XOS, which was most weighty at higher speeds (e.g. lower SFs).

(d) Cross-orientation enhancement and facilitation

An unexpected result was the widespread finding of small levels of XOF (typically less than 3 dB) for D.J.H., particularly at the slower stimulus speeds. T.S.M. also showed this effect, but only at the slowest speed tested (4 c deg^{-1} , 1 Hz). In a parallel study, we found XOF for superimposed cross-oriented masks for all the four observers tested at an SF of 7 c deg^{-1} and a stimulus pulse-duration of 400 ms (Meese & Holmes 2006; Meese *et al.* 2006). Work by others has found facilitation from pairs of cross-oriented flanking patches (Chen & Tyler 2002) and cross-oriented annuli (Yu *et al.* 2002). We compared the annulus effect with the superimposed effect and found them to be similar for a SF of 7 c deg^{-1} (Meese & Holmes 2006; Meese *et al.* 2006).

In the model here, the facilitatory influence (XOE) is constant across spatio-temporal frequency, but XOF is not seen at higher speeds where XOE is overpowered by masking. XOF/XOE is clearly more widespread than previously suspected, but what might be its cause? Chen & Tyler and Yu *et al.* attributed it to sensory interactions, and our equation (3.1) is consistent with that interpretation. For simplicity, we have supposed direct modulation by the mask mechanism on the detection mechanism. However, our results might also be understood in terms of disinhibition, involving an

intermediate process or mechanism. Another possibility is that the mask reduces uncertainty (Pelli 1985; Petrov *et al.* 2006). However, we think it unlikely that this is the sole cause of the facilitatory effects here. In our experiment, a quad of fixation points provided a strong positional cue throughout (cf. the thin ring used by Petrov *et al.* 2006), meaning that the mask provided a cue only to spatio-temporal frequency. As spatio-temporal conditions were blocked, the observers were well aware of what they were trying to detect on each trial. Furthermore, uncertainty cannot explain the effects of XOF that have been found on perceived contrast from (i) annular surrounds (Xing & Heeger 2001; Yu *et al.* 2001) and (ii) superimposed cross-oriented masks from a different SF band (Meese & Hess 2005, unpublished observations). And neither can it explain the single-cell evidence for XOF on response rate (Sillito & Jones 1996; Cavanaugh *et al.* 2002; Jones *et al.* 2002). Nevertheless, we cannot rule out the possibility that reduction of uncertainty made some contribution to the facilitatory effects seen here.

5. CONCLUSIONS

XOS is a well-established process in single-cell physiology (Morrone *et al.* 1982; Bonds 1989), visual-evoked potentials (Burr & Morrone 1987; Candy *et al.* 2001) and visual psychophysics (Ross & Speed 1991; Foley 1994). However, the tacit assumption of scale invariance, made by many contemporary modellers, is wrong. In human vision, XOS depends upon both SF and TF and poses a challenge to all the theories based on the scale-invariant properties of natural images (Field 1987). XOS resides in the high-speed corner of spatio-temporal space and to a first approximation declines in inverse proportion to the square root of stimulus speed (TF/SF). We suggest that this characteristic might reflect the system requirements of the underlying neurophysiology, where the contrast-invariant maintenance of population codes requires a suppressive gain pool in the fast and saturating

magnocellular stream, but less so in the slow and more linear parvocellular stream. The phenomenon of XOF is not restricted to lateral interactions as it had previously appeared, and the underlying process (XOE) is possibly ubiquitous for cross-oriented textures in close proximity on the retina. Both XOS and XOE depend upon mask contrast, but XOS grows much more rapidly than does XOE. On this view, the two processes of XOS and XOE are in competition across the entire spatio-temporal window, but the advantage (whether a mask produces XOM or XOF) depends upon spatio-temporal frequency and mask contrast. The functional significance of XOE remains to be elucidated.

This work was supported by a grant from the Engineering and Physical Sciences Research Council (GR/S74515/01) awarded to T.S.M. and Mark Georgeson. We would like to thank Mark Georgeson for his comments on an earlier draft.

ENDNOTE

¹In this paper, we use the terms cross-orientation masking (XOM) and cross-orientation facilitation (XOF) to refer to experimental phenomena: whether contrast detection thresholds are raised or lowered by a mask, respectively. The terms cross-orientation suppression (XOS) and cross-orientation enhancement (XOE) refer to the processes that underpin these phenomena. In our functional model, the parameter w is associated with XOS and the parameter α is associated with XOE. In some places, the terms XOM and XOS, and XOF and XOE can be used interchangeably.

REFERENCES

- Albrecht, D. G. & Geisler, W. S. 1991 Motion selectivity and the contrast-response function of simple cells in the visual cortex. *Vis. Neurosci.* **7**, 531–546.
- Allison, J. D., Smith, K. R. & Bonds, A. B. 2001 Temporal frequency tuning of cross-orientation suppression in the cat striate cortex. *Vis. Neurosci.* **18**, 941–948.
- Anderson, S. J. & Burr, D. C. 1989 Receptive field properties of human motion detector units inferred from spatial frequency masking. *Vis. Res.* **29**, 1343–1358. (doi:10.1016/0042-6989(89)90191-0)
- Baker, D. H. & Meese, T. S. 2006 Cross-orientation suppression occurs before binocular summation: evidence from masking and adaptation. *J. Vis.* **6**, 821a.
- Baker, D. H., Meese, T. S. & Summers, R. J. Submitted. Psychophysical evidence for two routes to suppression before binocular summation of signals in human vision.
- Barrett, B. T., McGraw, P. V. & Morrill, P. 2002 Perceived contrast following adaptation: the role of adapting stimulus visibility. *Spat. Vis.* **16**, 5–19. (doi:10.1163/15685680260433878)
- Bird, C. M., Henning, G. B. & Wichmann, F. A. 2002 Contrast discrimination with sinusoidal gratings of different spatial frequency. *J. Opt. Soc. Am. A* **19**, 1267–1273.
- Bishop, P. O., Coombs, J. S. & Henry, G. H. 1973 Receptive fields of simple cells in the cat striate cortex. *J. Physiol.* **231**, 31–60.
- Blakemore, C. & Campbell, F. W. 1969 On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. *J. Physiol.* **203**, 237–260.
- Bonds, A. B. 1989 Role of inhibition in the specification of orientation selectivity of cells in the cat striate cortex. *Vis. Neurosci.* **2**, 41–55.
- Bonds, A. B. 1991 Temporal dynamics of contrast gain in single cells in the cat striate cortex. *Vis. Neurosci.* **6**, 239–255.
- Bonin, V., Mante, V. & Carandini, M. 2005 The suppressive field of neurons in lateral geniculate nucleus. *J. Neurosci.* **25**, 10 844–10 856. (doi:10.1523/JNEUROSCI.3562-05.2005)
- Brady, N. & Field, D. J. 2000 Local contrast in natural images: normalisation and coding efficiency. *Perception* **29**, 1041–1055. (doi:10.1068/p2996)
- Burbeck, C. A. & Kelly, D. H. 1981 Contrast gain measurements and the transient/sustained dichotomy. *J. Opt. Soc. Am.* **71**, 1335–1342.
- Burr, D. C. & Morrone, M. C. 1987 Inhibitory interactions in the human vision system revealed in pattern-evoked potentials. *J. Physiol.* **389**, 1–21.
- Campbell, F. W. & Kulikowski, J. J. 1966 Orientational selectivity of the human visual system. *J. Physiol.* **187**, 437–445.
- Campbell, F. W. & Robson, J. G. 1968 Application of the Fourier analysis to visibility of gratings. *J. Physiol.* **197**, 551–566.
- Candy, T. R., Skoczenski, A. M. & Norcia, A. M. 2001 Normalization models applied to orientation masking in the human infant. *J. Neurosci.* **21**, 4530–4541.
- Carandini, M., Heeger, D. J. & Movshon, J. A. 1997 Linearity and normalization in simple cells of the macaque primary visual cortex. *J. Neurosci.* **17**, 8621–8644.
- Cavanaugh, J. R., Bair, W. & Movshon, J. A. 2002 Selectivity and spatial distribution of signals from the receptive field surround in macaque V1 neurons. *J. Neurophysiol.* **88**, 2547–2556. (doi:10.1152/jn.00693.2001)
- Chen, C.-C. & Tyler, C. W. 2002 Lateral modulation of contrast discrimination: Flanker orientation effects. *J. Vis.* **2**, 520–530.
- Clatworthy, P. L., Chirimuuta, M., Lauritzen, J. S. & Tolhurst, D. J. 2003 Coding of the contrasts in natural images by populations of neurons in primary visual cortex (V1). *Vis. Res.* **43**, 1983–2001. (doi:10.1016/S0042-6989(03)00277-3)
- Daugman, J. G. 1984 Spatial visual channels in the Fourier plane. *Vis. Res.* **24**, 891–910. (doi:10.1016/0042-6989(84)90065-8)
- DeAngelis, G. C., Robson, J. G., Ohzawa, I. & Freeman, R. D. 1992 Organization of suppression in receptive fields of neurons in cat visual cortex. *J. Neurophysiol.* **68**, 144–163.
- Derrington, A. M. & Lennie, P. 1984 Spatial and temporal contrast sensitivities of neurones in lateral geniculate nucleus of macaque. *J. Physiol.* **357**, 219–240.
- Ferrera, V. P. & Wilson, H. R. 1985 Spatial frequency tuning of transient non-oriented units. *Vis. Res.* **25**, 67–72. (doi:10.1016/0042-6989(85)90081-1)
- Field, D. J. 1987 Relations between the statistics of natural images and the response properties of cortical cells. *J. Opt. Soc. Am. A* **4**, 2379–2394.
- Field, D. J. & Wu, M. 2004 An attempt towards a unifying account of non-linearities in visual neurons. *J. Vis.* **4**, 283a.
- Foley, J. 1994 Human luminance pattern-vision mechanisms: masking experiments require a new model. *J. Opt. Soc. Am. A* **11**, 1710–1719.
- Foley, J. & Chen, C. 1997 Analysis of the effect of pattern adaptation on pattern pedestal effects: a two-process model. *Vis. Res.* **37**, 2779–2788. (doi:10.1016/S0042-6989(97)00081-3)
- Freeman, T., Durand, S., Kiper, D. & Carandini, M. 2002 Suppression without inhibition in visual cortex. *Neuron* **35**, 759–771. (doi:10.1016/S0896-6273(02)00819-X)

- Georgeson, M. A. & Meese, T. S. 2004 Contrast discrimination and pattern masking: contrast gain control with fixed additive noise. *Perception* **33**, 754–755.
- Graham, N. 1989 *Visual pattern analyzers*. Oxford, UK: Oxford University Press.
- Graham, N. & Sutter, A. 1998 Spatial summation in simple (Fourier) and complex (non-Fourier) texture channels. *Vis. Res.* **38**, 231–257. (doi:10.1016/S0042-6989(97)00154-5)
- Graham, N. & Sutter, A. 1996 Effect of spatial scale and background luminance on the intensive and spatial nonlinearities in texture segregation. *Vis. Res.* **36**, 1371–1390.
- Harvey, L. O. & Doan, V. Y. 1990 Visual masking at different polar angles in the two-dimensional Fourier plane. *J. Opt. Soc. Am. A* **7**, 116–126.
- Heeger, D. 1992 Normalization of cell responses in cat striate cortex. *Vis. Neurosci.* **9**, 181–197.
- Henning, G. B., Hertz, B. G. & Broadbent, D. E. 1975 Some experiments bearing on the hypothesis that the visual system analyses spatial patterns in independent bands of spatial frequency. *Vis. Res.* **15**, 887–897. (doi:10.1016/0042-6989(75)90228-X)
- Holmes, D. J. & Meese, T. S. 2004 Grating and plaid masks indicate linear summation in a contrast gain pool. *J. Vis.* **4**, 1080–1089. (doi:10.1167/4.12.7)
- Itti, L., Koch, C. & Braun, J. 2000 Revisiting spatial vision: toward a unifying model. *J. Opt. Soc. Am. A* **17**, 1899–1917.
- Jones, H. E., Wang, W. & Sillito, A. M. 2002 Spatial organization and magnitude of orientation contrast interactions in primate V1. *J. Neurophysiol.* **88**, 2796–2808. (doi:10.1152/jn.00403.2001)
- Kaplan, E. & Shapley, R. M. 1986 The primate retina contains two types of ganglion cells, with high and low contrast sensitivity. *Proc. Natl Acad. Sci. USA* **83**, 2755–2757. (doi:10.1073/pnas.83.8.2755)
- Legge, G. & Foley, J. 1980 Contrast masking in human vision. *J. Opt. Soc. Am. A* **70**, 1458–1471.
- Lennie, P. & Movshon, J. A. 2005 Coding of colour and form in the geniculostriate visual pathway (invited review). *J. Opt. Soc. Am. A* **22**, 2013–2033. (doi:10.1364/JOSAA.22.002013)
- Li, B., Peterson, M., Thompson, J., Duong, T. & Freeman, R. 2005 Cross-orientation suppression: monoptic and dichoptic mechanisms are different. *J. Neurophysiol.* **94**, 1645–1650. (doi:10.1152/jn.00203.2005)
- Meese, T. S. 2004 Area summation and masking. *J. Vis.* **4**, 930–943. (doi:10.1167/4.10.8)
- Meese, T. S. & Holmes, D. J. 2002 Adaptation and gain pool summation: alternative models and masking data. *Vis. Res.* **42**, 1113–1125. (doi:10.1016/S0042-6989(01)00291-7)
- Meese, T. S. & Holmes, D. J. 2003 Orientation masking: suppression and orientation bandwidth. *Perception* **32**, 388.
- Meese, T. S. & Holmes, D. J. 2006 Cross-orientation suppression is proportional to the square-root of speed for flickering Gabor stimuli. *J. Vis.* **6**, 200a.
- Meese, T. S., Holmes, D. J., Summers, R. J. & Wallis, S. A. 2006 Cross orientation suppression is not scale-invariant in space or time. *Perception* **35**, 417–418.
- Merigan, W. H. & Eskin, T. A. 1986 Spatio-temporal vision of macaques with severe loss of P beta retinal ganglion cells. *Vis. Res.* **26**, 1751–1761.
- Merigan, W. H., Byrne, C. E. & Maunsell, J. H. R. 1991 Does primate motion perception depend on the magnocellular pathway? *J. Neurosci.* **11**, 3422–3429.
- Morrone, M. C., Burr, D. C. & Maffei, L. 1982 Functional implications of cross-orientation inhibition of cortical visual cells I. Neurophysiological evidence. *Proc. R. Soc. B* **216**, 335–354.
- Morrone, M. C., Burr, D. C. & Ross, J. 1983 Added noise restores recognition of coarse quantized images. *Nature* **305**, 226–228. (doi:10.1038/305226a0)
- Nelson, S. B. 1991 Temporal interactions in the cat visual system. I. Orientation-selective suppression in the visual cortex. *J. Neurosci.* **11**, 344–356.
- Olshausen, B. A. & Field, D. J. 2005 How close are we to understanding V1? *Neural Comput.* **17**, 1665–1699.
- Olzak, L. A. 1985 Interactions between spatially tuned mechanisms: converging evidence. *J. Opt. Soc. Am. A* **2**, 1551–1559.
- Olzak, L. A. & Thomas, J. P. 1991 When orthogonal orientations are not processed independently. *Vis. Res.* **31**, 51–57. (doi:10.1016/0042-6989(91)90073-E)
- Olzak, L. A. & Thomas, J. P. 1999 Neural recoding in human pattern vision: model and mechanisms. *Vis. Res.* **39**, 231–256. (doi:10.1016/S0042-6989(98)00122-9)
- Pelli, D. G. 1985 Uncertainty explains many aspects of visual contrast detection and discrimination. *J. Opt. Soc. Am. A* **2**, 1508–1532.
- Pelli, D. G. 1987 On the relation between summation and facilitation. *Vis. Res.* **27**, 119–123. (doi:10.1016/0042-6989(87)90148-9)
- Petrov, Y., Carandini, M. & McKee, S. 2005 Two distinct mechanisms of suppression in human vision. *J. Neurosci.* **25**, 8704–8707. (doi:10.1523/JNEUROSCI.2871-05.2005)
- Petrov, Y., Verghese, P. & McKee, S. P. 2006 Collinear facilitation is largely uncertainty reduction. *J. Vis.* **6**, 170–178. (doi:10.1167/6.2.8)
- Phillips, G. C. & Wilson, H. R. 1984 Orientation bandwidths of spatial mechanisms measured by masking. *J. Opt. Soc. Am. A* **1**, 226–232.
- Rohaly, A. M., Ahumada, A. J. & Watson, A. B. 1997 Object detection in natural backgrounds predicted by discrimination performance and models. *Vis. Res.* **37**, 3225–3235. (doi:10.1016/S0042-6989(97)00156-9)
- Ross, J. & Speed, H. D. 1991 Contrast adaptation and contrast masking in human vision. *Proc. R. Soc. B* **246**, 61–69.
- Ross, J., Speed, H. D. & Morgan, M. J. 1993 The effects of adaptation and masking on incremental thresholds for contrast. *Vis. Res.* **33**, 2051–2056.
- Schwartz, O. & Simoncelli, E. P. 2001 Natural signal statistics and sensory gain control. *Nat. Neurosci.* **4**, 819–825. (doi:10.1038/90526)
- Sclar, G., Maunsell, J. H. R. & Lennie, P. 1990 Coding of image-contrast in central visual pathways of the macaque monkey. *Vis. Res.* **30**, 1–10. (doi:10.1016/0042-6989(90)90123-3)
- Sengpiel, F. & Vorobyov, V. 2005 Intracortical origins of interocular suppression in the visual cortex. *J. Neurosci.* **25**, 6394–6400. (doi:10.1523/JNEUROSCI.0862-05.2005)
- Seriès, P., Latham, P. E. & Pouget, A. 2004 Tuning curve sharpening for orientation selectivity: coding efficiency and the impact of correlations. *Nat. Neurosci.* **7**, 1129–1135.
- Sillito, A. M. 1979 Inhibitory mechanisms influencing complex cell orientation selectivity and their modification at high resting discharge levels. *J. Physiol.* **271**, 699–720.
- Sillito, A. M. & Jones, H. E. 1996 Context-dependent interactions and visual processing in V1. *J. Physiol. Paris* **90**, 2005–2209.
- Snowden, R. J. 1994 Adaptability of the visual-system is inversely related to its sensitivity. *J. Opt. Soc. Am. A* **11**, 25–32.
- Snowden, R. J. & Hammett, S. T. 1992 Subtractive and divisive adaptation in the human visual system. *Nature* **355**, 248–250. (doi:10.1038/355248a0)

- Solomon, S. G. & Lennie, P. 2005 Chromatic gain control in visual cortical neurons. *J. Neurosci.* **25**, 4779–4792. (doi:10.1523/JNEUROSCI.5316-04.2005)
- Somers, D. C., Nelson, S. B. & Sur, M. 1995 An emergent model of orientation selectivity in cat visual cortical simple cells. *J. Neurosci.* **15**, 5448–5465.
- Spear, P. D., Moore, R. J., Kim, C. B., Xue, J. T. & Tumosa, N. 1994 Effects of aging on the primate visual-system: spatial and temporal processing by lateral geniculate neurons in young adult and old rhesus monkeys. *J. Neurophysiol.* **72**, 402–420.
- Thomas, J. P. & Olzak, L. A. 1997 Contrast gain control and fine spatial discriminations. *J. Opt. Soc. Am. A* **14**, 2392–2404.
- Watson, A. B. & Solomon, J. A. 1997 A model of visual contrast gain control and pattern masking. *J. Opt. Soc. Am.* **14**, 2379–2391.
- Wilson, H. R., McFarlane, D. K. & Phillips, G. C. 1983 Spatial frequency tuning of orientation selective units estimated by oblique masking. *Vis. Res.* **23**, 873–882. (doi:10.1016/0042-6989(83)90055-X)
- Xing, J. & Heeger, D. J. 2001 Measurement and modeling of centre-surround suppression and enhancement. *Vis. Res.* **41**, 571–583. (doi:10.1016/S0042-6989(00)00270-4)
- Yu, C., Klein, A. S. & Levi, D. M. 2001 Surround modulation of perceived contrast and the role of brightness induction. *J. Vis.* **1**, 18–31. (doi:10.1167/1.1.3)
- Yu, C., Klein, S. A. & Levi, D. M. 2002 Facilitation of contrast detection by cross-oriented surround stimuli and its psychophysical mechanism. *J. Vis.* **2**, 243–255.
- Yu, C., Klein, S. A. & Levi, D. M. 2003 Cross- and iso-oriented surrounds modulate the contrast response function: the effect of surround contrast. *J. Vis.* **3**, 527–540. (doi:10.1167/3.8.1)