SUMMATION OF CONTRAST ACROSS THE VISUAL FIELD: A COMMON "FOURTH ROOT" RULE HOLDS FROM THE FOVEA TO THE PERIPHERY

A bioRχiv PREPRINT

Alex S Baldwin*
McGill Vision Research
Department of Ophthalmology & Visual Sciences
McGill University
Montreal, QC, Canada

Tim S Meese
College of Health and Life Sciences
Aston University
Birmingham, UK

May 6, 2025

ABSTRACT

Increasing the area of a grating stimulus tends to reduce the contrast at which it can be detected. We can use the relationship between stimulus area and this "threshold" contrast to study the detection process. Contrast signals combine linearly over short distances, as if summed within the receptive fields that define the inputs to early visual neurons. Beyond this range, the effects of area summation decrease. We set out to find whether the relationship between stimulus area and threshold is the same across the visual field. We measured thresholds for detecting "tiger tail" strips of grating (that grow orthogonal to the major axis of those presumed early receptive fields) in the fovea, parafovea (3 deg), and periphery (10.5 deg). The interpretation of previous studies taking similar approaches has been complicated by the variation in local contrast sensitivity across the visual field. In our study, we have mapped a detailed "Witch Hat" attenuation surface of this inhomogeneity for our three participants. We used this both in our model to account for the results, as well as to generate "compensated" stimuli that are equally detectable at each location within the stimulus extent. Our results follow a common fourth root summation rule in the fovea, parafovea, and periphery. Our "noisy energy" model predicts this behaviour through a combination of: i) the Witch Hat surface, ii) linear filtering by receptive fields that mimic those of V1 simple cells, iii) square-law contrast transduction, and iv) the application of an internal template that devotes the participant's attention to the extent of the stimulus. Fitting this model with one global sensitivity parameter (per participant) accounts for the foveal and parafoveal data (56 thresholds), with one further parameter allowing us to also model the periphery (84 thresholds).

Keywords vision · psychophysics · contrast sensitivity · summation · integration · visual field

*corresponding author: alexsbaldwin@gmail.com

1 Introduction

1.1 Summation of contrast to threshold for stimuli of different sizes

Area summation studies investigate how the visual system combines signals over space. For example, one can measure the relationship between the area of a sinusoidal grating and its contrast detection threshold (the lowest contrast at which a stimulus is detected with criterion probability). The Michelson contrast of a stimulus is defined as

$$C_{\text{Mich}} = \frac{L_{\text{max}} - L_{\text{min}}}{L_{\text{max}} + L_{\text{min}}}, \tag{1}$$

and is proportional to the square root of the power of the sinusoid that renders the grating stimulus.

Under Signal Detection Theory (SDT; Green and Swets, 1966), the ability to detect low-contrast targets is limited by intrinsic noise in the visual system (Gregory and Cane, 1955). In general, larger gratings can be detected at lower Michelson contrast. This produces a negative slope in a log-log plot of contrast threshold against stimulus area, where the steepness of the slope provides information about the summation process. In this article, we plot contrast thresholds expressed in logarithmic dB units on the y-axis

$$C_{\text{dB}} = 20 \times \log_{10}\left(C_{\text{Mich}}\right), \tag{2}$$

and, for ease of comparison, we apply the same transformation to the stimulus area for the x-axis.

A slope of -1 on a log-log threshold-vs-area plot is the signature of a linear combination of contrast across the visual field (provided the detection-limiting noise remains constant as the stimulus size increases). For linear summation, performance for detecting a small grating with a Michelson contrast of 2% would be the same as that for detecting a grating with twice the area at a contrast of 1%. Summation slopes that are shallower than this linear prediction may therefore indicate nonlinear characteristics of the detection process. For example, if the threshold were determined by the contrast energy of the stimulus (which is proportional to the contrast power integrated over space) then the summation slope would be -½ (Manahilov and Simpson, 1999).

1.2 The probability summation account of area summation to threshold

Empirical results from previous psychophysical studies have found summation slopes of approximately -¼ ("fourth root" summation; Bonneh and Sagi, 1999). These have often been interpreted to reflect probability summation (Robson and Graham, 1981), under which each part of a strip of grating is detected by independent local mechanisms. The number of possible detecting mechanisms therefore increases with the area of the grating. This results in greater sensitivity to larger stimuli, as there is a greater probability that at least one mechanism will detect the target (Sachs et al., 1971).

Probability summation models were originally formulated under High Threshold Theory (Green and Swets, 1966) where local mechanisms have a binary response ("detected" or not). A summation slope prediction (-1/β) can be derived from the slope (β) parameter of a Weibull psychometric function (Quick, 1974; Robson and Graham, 1981). The psychometric slope characterises how the probability for detection changes with signal strength. Under High Threshold Theory, when $\beta = 4$ the probability summation prediction corresponds to fourth root summation.

It has been demonstrated, however, that psychophysical behaviour is inconsistent with the binary High Threshold Theory concept (Swets, 1961a,b; Corso, 1963; Nachmias, 1981). Instead, results suggest that observers have access to a continuous response from each local mechanism, subject to additive noise. Consider an M-interval forced-choice task, where the observer must indicate in which of M temporal intervals a target was presented. Under SDT, if the target stimulates a single visual mechanism, such that the task is a decision over M response values, the decision can be made through a max operation where the single greatest response indicates the presence of the target (Tanner and Swets, 1954). This can be extended to develop a model of spatial summation. When multiple mechanisms are available to detect the target, a maximum response can also be taken over space in each interval. This means the decision is based on the single greatest activation that occurs in each trial. Under plausible assumptions, this model predicts fourth root summation (Tyler and Chen, 2000; Meese and Summers, 2012; Kingdom et al., 2015). For convenience, we will continue to refer to this as "probability summation", though the term may be a misnomer under SDT.

1.3 Evidence for a noisy energy model

Our recent work on area summation has favoured a "noisy energy" model over probability summation (Meese and Summers, 2007; Meese, 2010; Meese and Summers, 2012; Baldwin and Meese, 2015). In this model, the responses of local contrast-driven mechanisms are squared and summed within a template that roughly corresponds with the size and location of the stimulus. The size of the template plays a double role in influencing threshold; not only does it set the integration range for the signal, but the response at each location is also perturbed by independent local noise. The variance of the noise that is integrated therefore increases in proportion to the area of the template.

In cases where the template in the noisy energy model exactly matches the stimulus area, summation obeys a fourth root rule (Meese and Summers, 2012). This results from a combination of: i) the square-law transduction of contrast into local mechanism responses, and ii) the quadratic effect of the integration of area-dependent noise and signal. For more complex stimuli the template-matching process breaks down (Meese and Summers, 2007; Baker and Meese, 2011; Meese and Summers, 2012; Baldwin and Meese, 2015). For example, an experiment using a "Battenberg" checkerboard pattern of signal regions (containing contrast) and non-signal regions (which were

blank) revealed square root (quadratic) summation (Meese, 2010). This can be explained by limitations in the template. Consider the case where a template can be matched in its extent to a square or circular grating, but cannot vary in the weightings applied within that region to account for the non-signal "gaps" in the Battenberg stimulus. The template is then matched to the expected outline of the stimulus rather than the arrangement of signal and non-signal pattern within it. The level of additive internal noise integrated by the template would then be constant, regardless of whether all or half of its internal area contained signal. In the noisy energy model, this knocks out the template effect and leaves only the quadratic effect of square-law signal transduction. This result therefore distinguishes the predictions of the noisy energy model from those of the probability summation models that produce fourth root rules (or thereabouts), as described above.

1.4 Effects of the visual field inhomogeneity in local contrast sensitivity

Most area summation experiments fall into one of two categories. In the first, contrast detection thresholds are measured for centrally-presented gratings for a wide range of grating areas (e.g. Howell and Hess, 1978; Rovamo et al., 1994; Meese and Summers, 2012). The summation slopes in these experiments are however confounded by the inhomogeneity in contrast sensitivity across the visual field (Hilz and Cavonius, 1974; Pointer and Hess, 1989). As the size of the grating increases it extends into more eccentric regions of the visual field where the sensitivity is lower. The variation in sensitivity across the visual field was mapped by Baldwin et al. (2012), who found the decline of local contrast sensitivity with eccentricity to be bi-linear on a log-log plot: initially steep and then shallow. These measurements were the basis for personalised "attenuation surfaces" of contrast sensitivity used by Baldwin and Meese (2015) to counteract the inhomogeneity by multiplying gratings of various sizes by the inverse of the surface. This allowed a cleaner measure of area summation by factoring out the confounding influence of the inhomogeneity in local sensitivity.

In the second type of study, smaller strips of grating are used to test local summation behaviour and the effects of stimulus aspect ratio (Robson and Graham, 1981; Mayer and Tyler, 1986; Polat and Norcia, 1998; Manahilov et al., 2001; Foley et al., 2007; Meese and Hess, 2007). These strips of grating are often presented to regions of the visual field outside the fovea where contrast sensitivity is more uniform. These strips can "grow" in size in a direction either parallel ("skunk tails" in Meese and Hess, 2007) or perpendicular ("tiger tails") to the bars of the grating. There is evidence for summation along skunk tail stimuli involving additional mechanisms (Chen and Tyler, 1999; Chen et al., 2019, 2023), perhaps related to the integration of contours (Field et al., 1993). In this study we are primarily concerned with summation along the width of our tiger tail stimuli (orthogonal to the grating stripes), though our use of stimuli with different heights does allow some examination of summation parallel to the grating stripes.

For very small stimuli with an area of less than one square cycle (Foley et al., 2007) the summation slopes are steep. This is thought to reflect linear summation within the footprint of early receptive fields (Chen and Tyler, 1999; Meese, 2010). Slightly larger stimuli presented to the periphery, with an area of less than 32 square cycles, have log-log summation slopes around -½ (Manahilov et al., 2001; Meese and Hess, 2007). For even larger stimuli, the slopes are around -¼ (Robson and Graham, 1981). Summation slopes for foveal stimuli are flatter due to the effect of the inhomogeneity in contrast sensitivity across the visual field. As the stimulus "grows" in size, it encroaches into progressively less sensitive regions of the visual field, meaning there is a diminishing improvement in detectability as the signal area increases (Baldwin et al., 2012).

1.5 Does area summation obey a single rule across the visual field?

These previous studies typically examine summation over only a relatively narrow range of stimulus sizes. In many cases, the stimuli were presented to a single visual field locus. No previous study has applied a direct compensation to counteract the inhomogeneity in contrast sensitivity. The summation slopes reported by these studies may not generalise beyond the relatively limited conditions under which they were measured. For example, the tendency for studies with smaller stimuli to report steeper summation slopes.

Our goal in the current study was to instead measure summation slopes for a wide range of stimulus sizes, at *three* visual field locations, both with and without compensation for visual field inhomogeneity (Baldwin and Meese, 2015). This allowed us to investigate whether the results could be accounted for by a common function. We present new results here for stimulus sizes that span the interesting transition between steep and shallow slopes that otherwise appear to divide the conclusions across previous studies. Our "Witch Hat" compensation for the inhomogeneity in contrast sensitivity flattens the effective contrast of foveal and parafoveal stimuli. Measurements with these stimuli are comparable to those made previously in the periphery where contrast sensitivity is relatively constant. The consistent behaviour we find once the inhomogeneity in sensitivity is factored out allows us to present a single model that accounts for summation across the visual field.

2 Methods

2.1 Participants

All participants were volunteers who gave their informed consent before participating in the study, which was conducted in accordance with the Declaration of Helsinki. Procedures were approved by the Ethics Committee of the School of Life and Health Sciences at Aston University (approval #856). The three participants (ASB, DHB and SAW) were 22, 28 and 46 years old respectively.

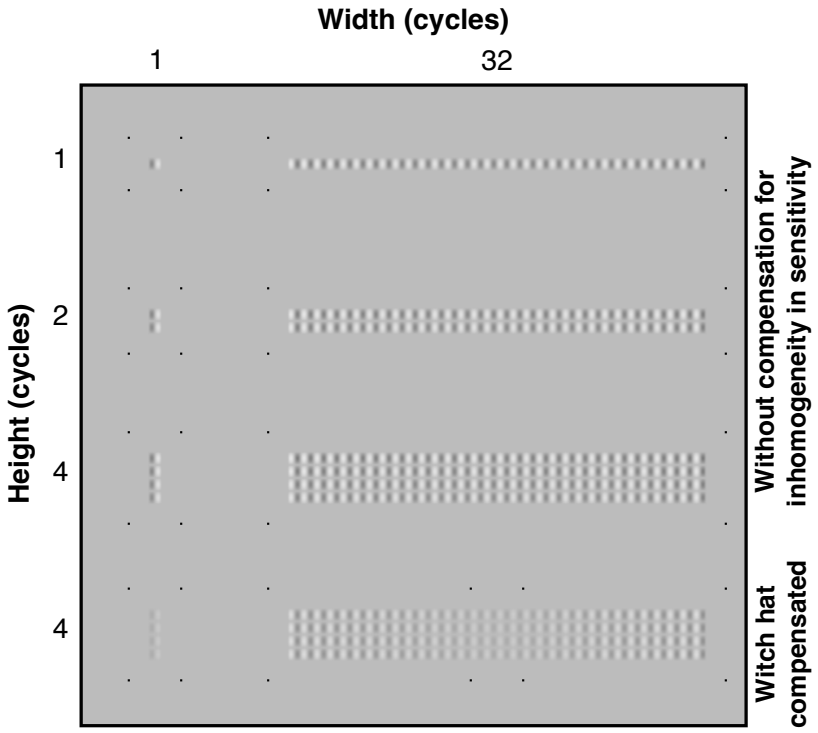

Fig 1: The first three rows show the narrowest (1 cycle, left) and widest (32 cycles, right) stimuli for all three heights (1, 2 and 4 cycles). These are the non-compensated stimuli, with each micro-pattern element rendered at the same contrast. The last row shows the effect of the Witch Hat compensation. Also shown are the quads of dots that served to indicate the stimulus extent to the participant.

Participants wore optical correction appropriate for the viewing distances tested when required. All experiments were performed binocularly with natural pupils.

2.2 Equipment

Stimulus presentation was performed using a CRS ViSaGe (Cambridge Research Systems, Rochester, UK), which was used with a gamma-corrected CRT monitor (Eizo Flexscan T68) to achieve 14-bit grayscale precision. The monitor had a refresh rate of 120 Hz, and a mean luminance of 75 cd/m^2. The monitor was viewed from a distance of 1.2 metres. At this distance, there were 48 pixels per degree of visual angle (deg), giving the 4 c/deg stimuli used in this study 12 pixels per cycle.

2.3 Stimuli

Our "tiger-tail" stimuli were formed from 4 c/deg vertical "Battenberg" signal elements (Meese, 2010) arranged into rectangular strips (Fig. 1). Each element was a single cycle of sinusoidal grating multiplied by an orthogonal cosine-phase half-cycle at half the target spatial frequency. An

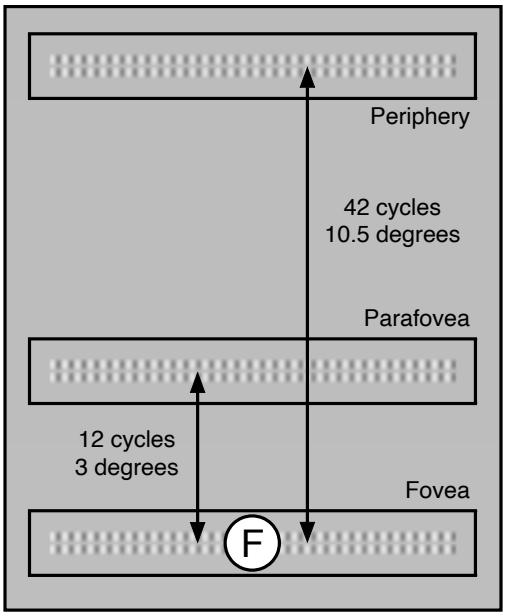

Fig 2: Stimulus locations used in this study (not to scale) for example stimuli two elements tall and thirty-two elements wide. The "F" icon marks the fixation location, but neither this nor a central fixation point was present in the study. Instead, the point of fixation was inferred from a quad of points around the stimulus in the foveal condition or marked directly by a 2 pixel square black dot in the parafoveal condition, or by a red LED in the peripheral condition.

alternative stimulus design would have been to define a family of rectangular gratings of various sizes. It is generally preferable to soften the edges of grating stimuli by modulating the contrast with an envelope such as a Gaussian (to make a Gabor patch) or a raised-cosine (as in Baldwin and Meese, 2015). This complicates the definition of the stimulus extent, and consequently the calculation of the effective stimulus area. For the Battenberg stimuli used here, each micropattern element contains its own envelope. This means that the area of a 1×4 rectangular stimulus is simply four times the area of a 1×1 stimulus. Our choice of micropattern stimuli over simple gratings made no difference to the sinusoidal horizontal cross-section taken through the centre of a "row" of elements: The abutting single cycles of sinusoidal grating combined into a single continuous sinusoid. In their vertical cross-section however, the envelopes of our stimuli were rectified cosines at half the spatial frequency of the stimulus. This is accommodated in our modelling.

Stimuli were presented with six different widths (1, 2, 4, 8, 16 and 32 cycles) and three different heights (equivalent to 1, 2 and 4 cycles). The 1 and 4 cycle high stimuli were used in all six stimulus width conditions, whereas the 2 cycle high stimuli were used only with the narrowest (1 cycle) and widest (32 cycles) stimulus widths. The nominal areas of the stimuli were calculated as the product of the width and height. These sizes were chosen to cover most of the range tested in previous studies. Our stimulus selection included a wider range of widths than heights, focusing our

study to investigate "tiger tail" summation along the axis perpendicular to the stripes of the carrier grating. As discussed above, we expected this summation may be simpler than that found for stimuli that increase in height, growing *along* the axis of their stripes ("skunk tails"). However, once any particular "skunk tail" summation effects were accounted for in a narrow stimulus (1 cycle wide) of a particular height (whether 1, 2, or 4 cycles), the possibility of a common "tiger tail" summation behaviour could then be investigated as a stimulus of that height grows in its width. Including stimuli of multiple heights and widths also allowed us to compare sensitivity to stimuli that have the same area but different combinations of widths and height (e.g. an area of four square cycles achieved with stimuli that are either one cycle in height and four cycles in width, or vice-versa). By including stimuli of different heights (extensions parallel to the grating orientation), we can measure sensitivity to sets of stimuli that have the same area but different aspect ratios.

Stimuli were generated both with flat contrast profiles and with Witch Hat compensation for the visual field inhomogeneity. The compensated stimuli were produced by multiplying the original stimuli by the inverse of the Witch Hat attenuation surface measured for each participant in Baldwin et al. (2012). This was to give them an effectively flat internal signal profile after the attenuating effects of visual field inhomogeneity (Baldwin and Meese, 2015). The nominal contrast of the compensated stimuli was their contrast before the compensation was applied. This means that the expected contrast threshold for a stimulus would no longer depend on its location in the visual field.

2.4 Procedures

The experiment was controlled using an in-house software suite ("Liberator") written in Delphi (Borland Software Corporation, California) at Aston University. Three eccentricity conditions were tested. Stimuli were presented centred either at central fixation or at two eccentric locations along the upper vertical meridian (either 12 or 42 stimulus carrier cycles superior to the fixation point, equivalent to 3 and 10.5 degrees of visual angle respectively), as shown in Fig. 2. We refer to these three field conditions as foveal (taken as the area within 2 deg of fixation), parafoveal (2 - 10 deg), and peripheral (beyond 10 deg), respectively. Because the stimuli extended over space, there was overlap in the eccentricities tested across the three field conditions. For the 32-cycle wide stimulus centred at fixation, the edges were at 4 deg. For parafoveal and peripheral stimuli the edges of the widest stimuli reached 5 deg and 11.2 deg respectively. Participants ASB and DHB were tested at all three eccentricities, whereas SAW was tested only for the foveal and parafoveal conditions.

The spatial extent of each stimulus was indicated by a quad of black dots surrounding its corners. For the foveal and parafoveal conditions, these dots were 2 x 2 pixel squares. In the peripheral condition, their size was increased to 4 x 4 pixels to ensure visibility. The method for marking the intended point of fixation depended on the field condition. For stimuli presented in the fovea, this

was an additional quad of dots around that location (Summers and Meese, 2009). In the parafoveal condition, we used a single 2-pixel square dot. In the peripheral condition, a dim red LED was used for fixation. This was positioned below the monitor, coplanar with the display screen, such that the distance between the LED and the centre of the stimulus was 10.5 degrees of visual angle.

Stimuli were blocked by size and location. The non-compensated and Witch Hat compensated conditions were interleaved to encourage participants to adopt the same attention strategy across the two conditions. This was motivated by our expectation that the wider non-compensated stimuli would be most visible at their centre. If this condition was tested in a separate block from the compensated stimuli (which should be equally visible across their entire span), the participant may have restricted their attentional window (Tyler and Chen, 2000) to the most salient region. The interleaving was designed to counter this potential confound. Thresholds were measured using a two-interval forced-choice three-down, one-up staircase procedure. Participants were given audible feedback as to whether they correctly chose the target interval. Each condition was repeated four times by each participant, testing the blocks in a randomised order. Contrast detection thresholds for each repetition were calculated by fitting a Weibull function to the data using Palamedes (Prins and Kingdom, 2018). Mean thresholds and standard errors (in dB) were calculated across repetitions.

3 Results

We first consider the results for different stimulus sizes and eccentricities before showing how the foveal and parafoveal data (56 thresholds per participant) can be modelled with a single fitted parameter per participant. Extending that to include the peripheral data (bringing us to 84 thresholds, and taking us beyond the region where we measured the Witch Hat attenuation surfaces used in our stimuli and modelling) requires one further fitted parameter per participant.

3.1 Width summation without compensation for visual field inhomogeneity

Contrast detection thresholds for stimuli of various widths presented to the fovea, parafovea, and periphery are shown in Fig 3. The subplots A-C show results from each of the three participants. Each data series shows a reduction in threshold as stimulus width increased but height remained constant at either 1 or 4 cycles (see legend). More generally, thresholds decreased with increasing stimulus area across conditions. For stimuli presented to the fovea (filled symbols), summation was initially at its steepest, and then flattened out for larger stimulus sizes. We attribute this to the visual field inhomogeneity for contrast sensitivity, consistent with previous foveal results (e.g. Robson and Graham, 1981; Polat and Tyler, 1999).

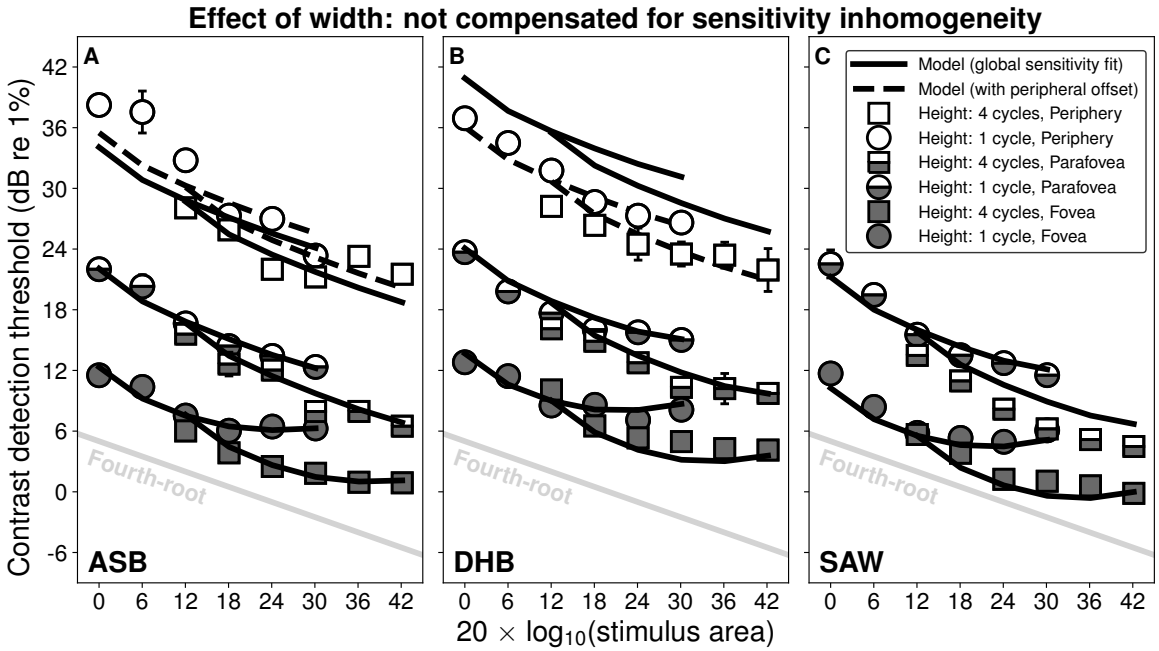

Fig 3: Area summation for stimuli with increasing width (lateral extent across the display). The subplots present data from three participants for uncompensated stimuli presented in the periphery (open), parafovea (half-filled), and fovea (filled symbols). Results are shown for stimuli that are 1 (circles) and 4 cycles (squares) in height. The black solid and dashed lines are thresholds from our model. The grey line shows a "fourth root" summation slope.

In the parafovea and periphery (half-open and open symbols in Fig 3), the benefit from increasing the size of the stimulus extended to larger stimulus areas; the decline in threshold does not flatten out to the same extent as in the foveal condition. This is consistent with previous investigations of summation away from the fovea where stimuli grow along a circular arc centred on fixation (maintaining the same eccentricity) or at a tangent to such an arc (maintaining a *similar* eccentricity). This largely avoids the confounding effects of the variation in sensitivity across the visual field (e.g. Robson and Graham, 1981; Mayer and Tyler, 1986; Manahilov et al., 2001).

3.2 Width summation with Witch Hat compensation

Applying the Witch Hat compensation to our stimuli counteracts the inhomogeneity in local contrast sensitivity (Baldwin et al., 2012), causing summation slopes to "straighten" (Fig 4). This reveals a greater consistency of contrast summation (approximately fourth root) spreading out from the fovea than was apparent with the non-compensated stimuli (those data replotted with pale symbols in Fig 4 for comparison). We have previously demonstrated a similar result with circular grating stimuli (Baldwin and Meese, 2015). Summation had much the same form across the full 32 cycles of the 4 c/deg stimuli, equivalent to 8 degrees of visual angle.

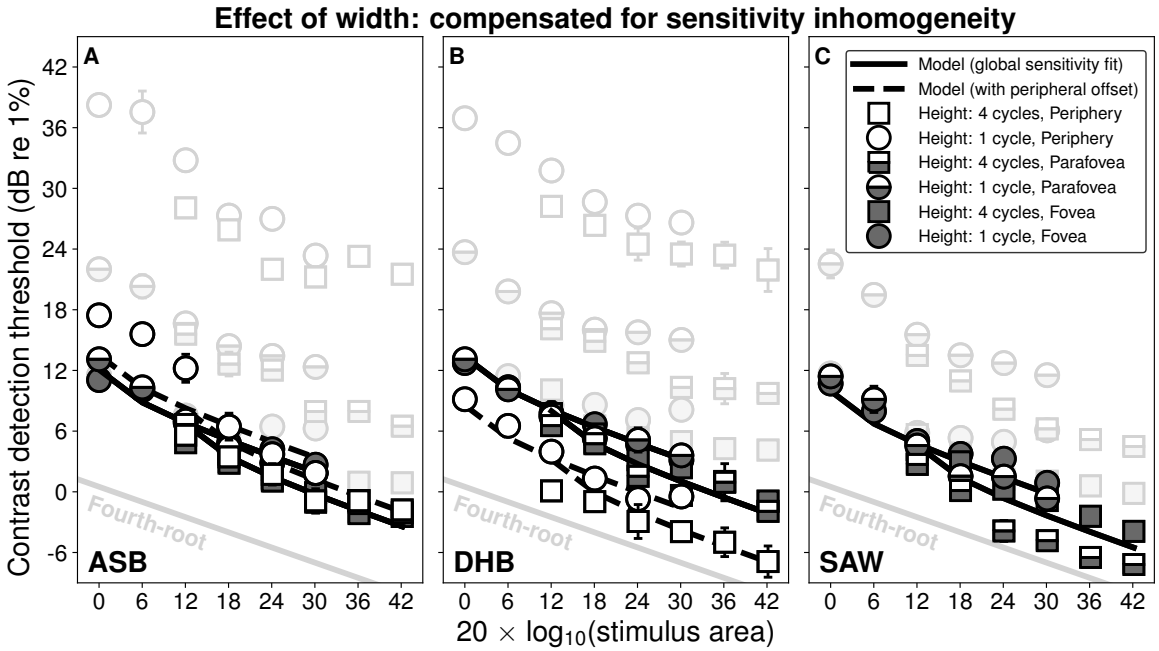

Fig 4: Results for comparable conditions to those in Fig 3, but where local stimulus contrast was adjusted to compensate for visual field inhomogeneity according to the "Witch Hat" of Baldwin et al. (2012). Results for the uncompensated condition (from Fig 3) are replotted here in pale grey. As compensation leads to the overlapping of results from the different visual field locations, we present results separately across field location in Fig 7.

The Witch Hat compensation had less effect on the summation curves for the parafoveal and peripheral stimuli due to the relative homogeneity of contrast sensitivity in this region. The results for both stimulus heights are consistent with approximately fourth root summation over 32 cycles in the parafovea, a slope shallower than that found by Manahilov et al. (2001) at this eccentricity.

Applying the Witch Hat compensation tended to equate thresholds obtained in the fovea and parafovea for all three participants. In Fig 4, this causes the half-filled (parafovea) symbols to overlap the filled symbols (fovea). This suggests that i) the attenuation surfaces we applied to our stimuli were accurate, and ii) summation in the fovea and parafovea follow the same rule. For participants ASB and DHB, this superposition was striking (Fig 4A-B). For SAW, there was some disparity between the results across the two field positions, the benefit of area for the 4 cycle high stimuli being greater in the parafovea than in the fovea (compare half open squares with filled squares in Fig 4C). We shall return to this point in the **Discussion**.

In the peripheral condition, the stimuli were presented to regions beyond where the attenuation surfaces were mapped in Baldwin et al. (2012). In that study, we measured eccentricities up to 18 cycles, less than half of the distance to the peripheral stimuli in the study here (42 cycles from the fovea at their closest point). We therefore rely on extrapolating the second (shallower) limb of

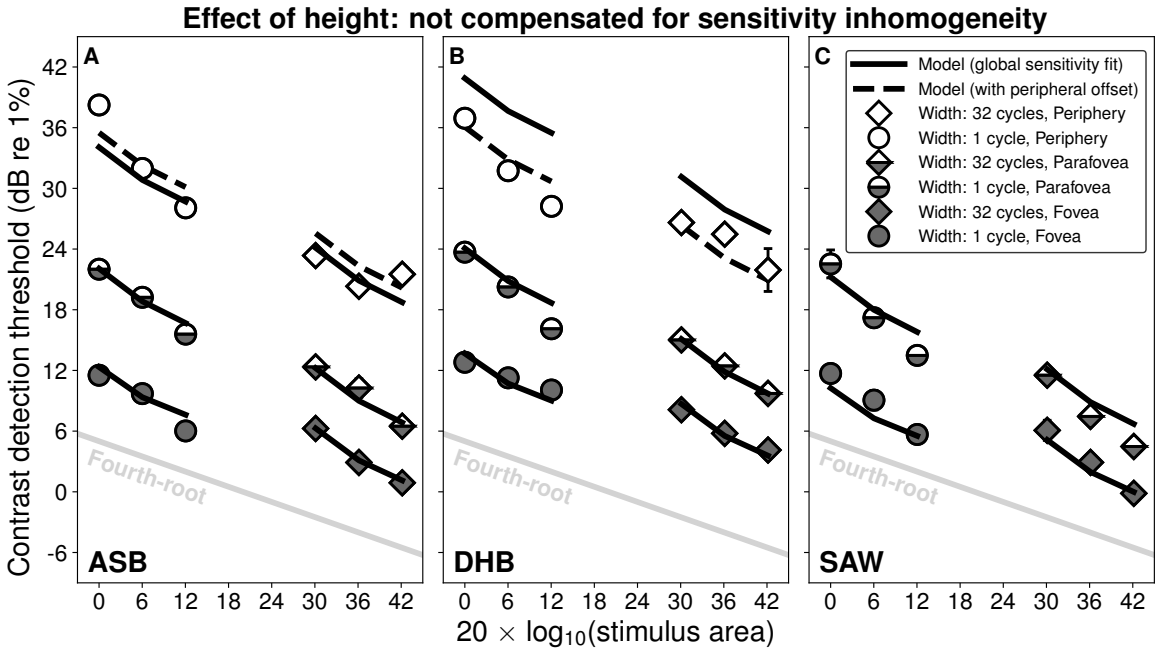

Fig 5: Area summation as functions of height (see the caption for Fig 3 for further details on the presentation). Results are shown for stimuli that are 1 cycle (circles) and 32 cycles (diamonds) in width. The first and last datapoint of each triplet (corresponding to heights of 1 and 4 cycles) are replotted from Fig 3. The intermediate point in each triplet is for a height of 2 cycles.

the bilinear decline from our Witch Hat model into the periphery. Nevertheless, the compensation for ASB did a good job of equating sensitivity (the open symbols tend to superimpose with the other symbols in Fig 4A), though there was a tendency for the smaller 1 cycle high stimuli to have higher thresholds than expected. For DHB (Fig 4B) the peripheral compensation overshot, resulting in lower nominal thresholds than for other stimuli of the same size. This implies a shortcoming from our simple extrapolation. In fact, additional measurements at greater eccentricities in Baldwin (2013a) suggest that the second limb of the Witch Hat surface's decline might be slightly shallower than estimated in Baldwin et al. (2012), consistent with the over-compensation for DHB here.

3.3 Height summation without compensation for visual field inhomogeneity

Results for the narrowest and widest stimuli from Fig 3 are replotted in Fig 5, with the addition of results for a stimulus height of 2 cycles. These plots show the effect of increasing the height of our narrowest and widest stimuli (triplets of data points). In comparison to Fig 3, which provides a direct illustration of summation steeper than "fourth root" only for the smallest sizes, we see a greater tendency for summation to outpace the fourth root rule when stimuli grow along the axis parallel to their stripes.

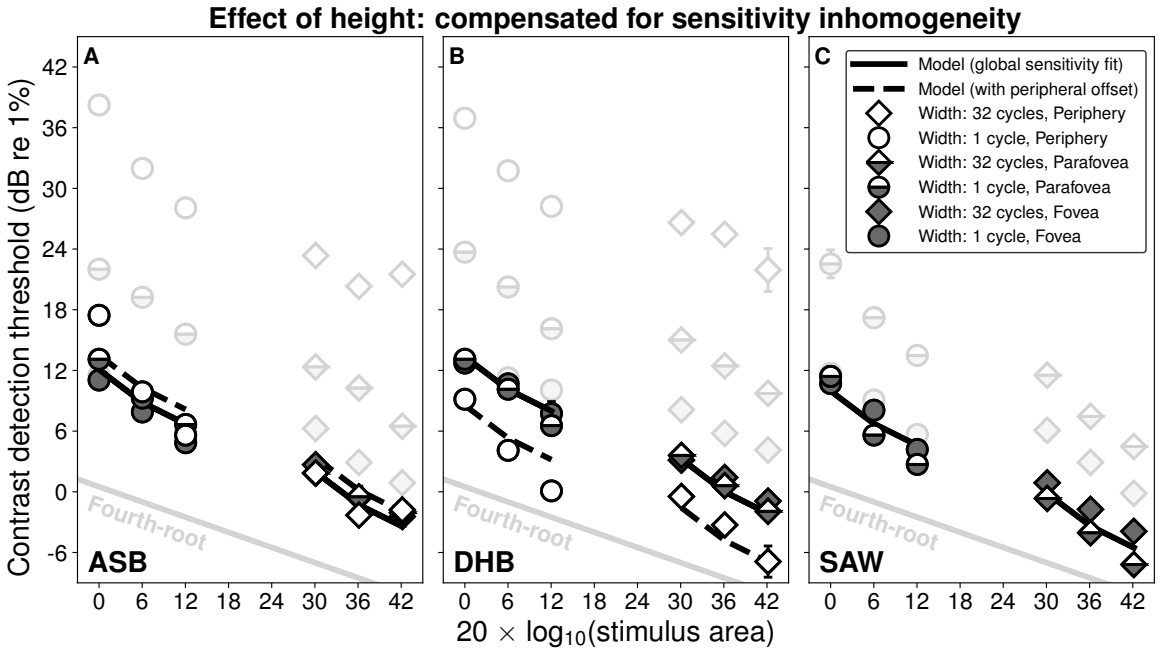

Fig 6: Area summation as functions of height where local stimulus contrast was varied to compensate for visual field inhomogeneity (see the caption for Fig 5 for further information on the format).

3.4 Height summation with Witch Hat compensation

With Witch Hat compensation applied to our stimuli of different heights (Fig 6), the results from the three visual field locations collapse as before (Fig 4). This also results in similar summation slopes in the fovea, parafovea, and periphery, though the effect of overcompensation from the extrapolated attenuation surfaces is seen again for DHB (Fig 6B; open symbols).

3.5 Modelling our spatial summation results

The results in Figs. 3-6 are shown with the predictions of the noisy energy model (Meese, 2010; Meese and Summers, 2007, 2012) with the form used in Baldwin and Meese (2015). The model predicts how detection thresholds should vary for stimuli of different sizes and locations, by finding the contrast that would result in a specific signal-to-noise ratio ($d' = 1$).

To calculate the numerator of the signal-to-noise ratio (the "signal" magnitude for a specific stimulus condition), the stimuli are first attenuated by the Witch Hat surface that was mapped previously for each participant (Baldwin et al., 2012). They are then filtered by Cartesian-separable log-Gabor wavelets (Baker et al., 2022) matched in spatial frequency and orientation to the stimuli. As filters, these wavelets have a spatial frequency bandwidth of 1.6 octaves and orientation bandwidth of 25°. They model the responses of neurones with receptive fields analogous to those of simple cells in primary visual cortex (De Valois et al., 1982). Within the footprint of these filter

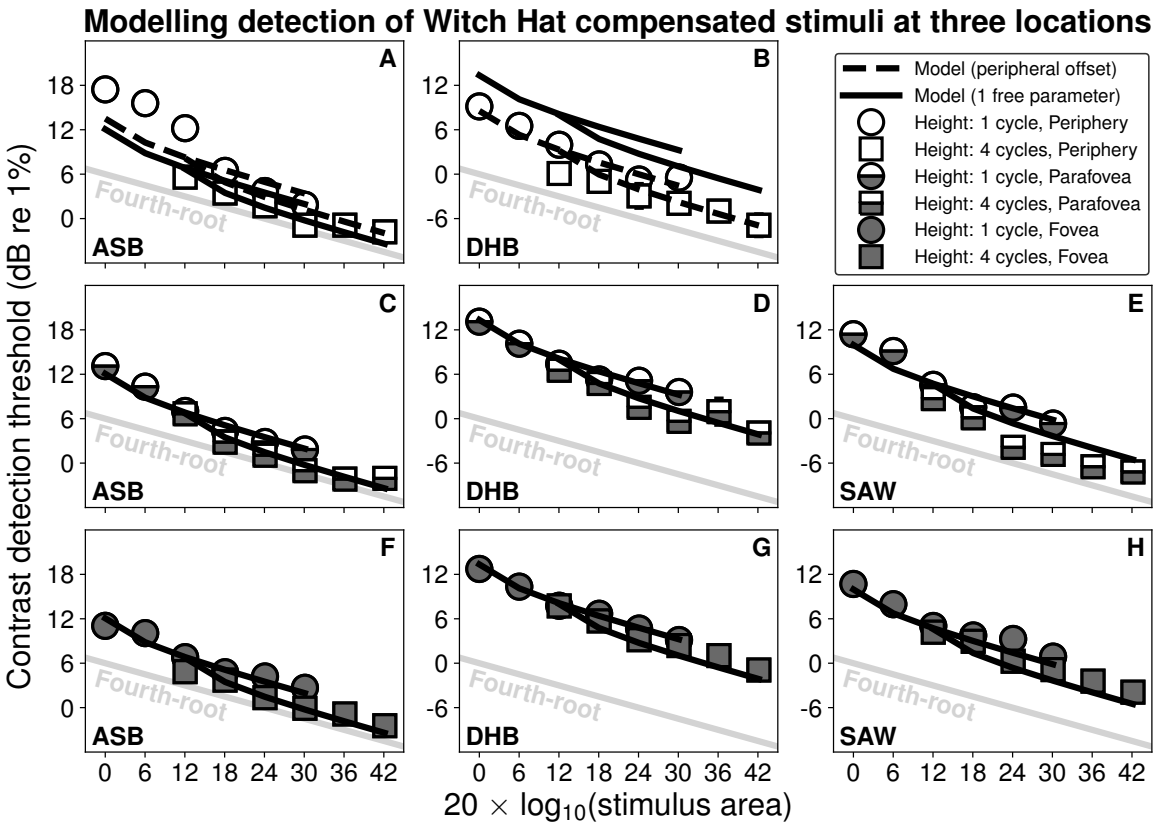

Fig 7: Results for Witch Hat compensated stimuli replotted from Fig 4. Separate panels are for the different field conditions to facilitate comparisons between human data and model predictions at each eccentricity.

elements, the responses to matching features in the stimulus will sum linearly. The outputs of these local filter-elements are squared (Meese, 2010) and then summed within a template that is matched in size to the outline of the stimuli. The template is not matched to the decline in contrast sensitivity across the visual field. For non-compensated stimuli, the ideal observer would attribute more weight to the regions of the visual field where the sensitivity is highest, however this small detail makes little difference to the model predictions for the stimulus and task design in the current study (Baldwin, 2013b). The square-law transduction of local contrasts in the model has the effect of reducing the summation slope to square root beyond the short-range of linear summation that occurs "within-filter".

The denominator of the signal-to-noise ratio is the standard deviation of the combined noise affecting the decision. We assume this originates from independent Gaussian noise sources affecting the response of each filter output. The larger template sizes include a greater number of these noise sources, with the standard deviation of the combined noise at the decision stage increasing with the square root of the template area. This approach assumes that the observer restricts the template size

Table 7: Global sensitivity parameters and RMS errors for the model fits to the results presented in Figs 3-7. A global sensitivity parameter fits the data measured in the fovea and parafovea as described in the text. The offset parameter in the periphery is a second free parameter to adjust sensitivity to the peripheral stimuli alone.

	Fovea and parafovea		**Periphery**	
Participant	**Global sensitivity (dB)**	**RMSe (dB)**	**Peripheral offset (dB)**	**RMSe (dB)**
ASB	11.3	0.8	1.4	2.4
DHB	12.6	0.9	-4.8	1.2
SAW	9.2	1.5		

to the stimulus size within an experimental block, thereby excluding the noise from irrelevant signal locations. By comparison to a fixed summation window (matched to the largest stimulus) this ideal approach (e.g. Tyler and Chen, 2000) reduces the model summation slope from linear to square root. However, with our square-law transduction of signal in place (see above) the combined effects reduce the model's summation slope from square root to fourth root.

To fit the model to foveal and parafoveal results, we adjusted a single global sensitivity parameter (an offset) for each participant (i.e., the model has one degree of freedom for each participant). This is equivalent to adjusting the common standard deviation of local additive noise sources affecting the response of each filter element. This also accommodates what might appear to be a discrepancy between our approaches to the data analysis and modelling. In our data analysis, the fitting of Weibull psychometric functions delivered "thresholds" for 82% correct, whereas the model predicted thresholds for a d' of unity, which is 76% correct (in our two-interval forced-choice task). Provided the slope of the psychometric function is invariant with stimulus area (a result found in several previous studies, including Mayer and Tyler, 1986; Meese and Summers, 2012; Wallis et al., 2013; Baldwin and Meese, 2015), the model adjustment needed to predict thresholds at 82% correct would be a constant offset (a lateral shift of the psychometric function). This effect is therefore absorbed into our fitted global sensitivity parameter.

For each participant, fits were made to all of the foveal and parafoveal thresholds using the *fminsearch* method in Matlab 2016a (Mathworks, Natick, MA). The fitting procedure found the parameter value that produced the lowest RMS error between the model predictions and the human data. The fitted parameter values and RMS errors are shown in Table 1.

Referring back to the model curves for the foveal and parafoveal locations in Fig 3 (solid black lines), the model (with a single fitted parameter) provides a good account of the results for all three participants. It predicts the initial steep (square root) fall in threshold where summation is accelerated by within-filter linear effects, the subsequent fourth root summation, and the differences in sensitivity between the fovea and the parafovea. The model also captures the shallowing of

the summation slope caused by the visual field inhomogeneity for the non-compensated stimuli. In Fig 5, we see how the same model predicts the steeper summation slopes found by growing our stimuli in height (parallel to their stripes) with no further parameters. This may be a simple consequence of linear summation within the simulated "receptive fields" in our model, which are elongated in an orientation parallel to the grating stripes. However, it is notable that effects of summation along stimuli with a width of one cycle in the parafovea or periphery outpaces the predictions of our model (which includes this within-filter linear summation) in all three participants. This may point to the influence of specialised mechanisms involved in this type of summation (Chen and Tyler, 1999; Chen et al., 2019, 2023).

For the Witch Hat compensated stimuli in Fig 4, the same model (with the same global sensitivity parameter) predicts that the results for the different field positions should be identical, because the effects of the visual field inhomogeneity have been factored out. The data and model fits are replotted separately for each field position in Fig 7 to facilitate visual assessment. The empirical results from the fovea and parafovea agree with this prediction with one exception. Participant SAW (Fig 4C) had lower thresholds for the 4 cycle high stimuli in the parafovea than predicted. We consider possible reasons for this in the **Discussion** below.

Results from two participants (ASB and DHB) for the peripheral condition are also shown in Figs 3-7. Within this region, contrast sensitivity is relatively constant (Robson and Graham, 1981), and so Witch Hat compensation has little effect on the shape of the summation slope. In general, both participants showed fourth root summation as the stimulus size increased up to its maximum width of 32 cycles, consistent with the results reported by Robson and Graham (1981) for this location. However, as mentioned above, the Witch Hat for the periphery is a long extrapolation from the surfaces measured in Baldwin et al. (2012). For ASB, this is quite successful, nonetheless, while for DHB our extrapolated attenuation surface predicted sensitivity to be slightly worse in the periphery than was found. To account for this, we re-fit these data with an additional "offset" parameter provided in Table 1. For ASB the optimal fit required an additional offset of only 1.4 dB since the initial extrapolation was quite close in any case. For DHB rather more correction was required (-4.8 dB).

A final discrepancy between human and model is for ASB's smallest stimuli. Both with (Fig 7) and without compensation (Fig 3), thresholds were higher than predicted in the periphery for stimuli that were one cycle high and up to four cycles wide. We propose an explanation for this in the **Discussion** below.

4 Discussion

4.1 Fourth root summation of contrast to threshold across the visual field

We have shown that the spatial summation of contrast to detection threshold can be explained by a common process in the fovea, parafovea, and periphery. The results are consistent with short-range linear summation within the receptive fields of local detecting mechanisms, followed by fourth root summation across them. We were able to reveal this equivalence by applying our Witch Hat attenuation surface (Baldwin et al., 2012), both as a component in our model and in the rendering of stimuli to compensate for local variations in contrast sensitivity.

In the periphery, our finding that summation for stimuli of more than a few cycles follows a fourth root rule agrees with several previous studies (Robson and Graham, 1981; Mayer and Tyler, 1986, referring to the shallower slope for the larger sizes in the latter study). For smaller stimuli (one to two cycles), we see the effects of linear summation within the model's filter elements (and presumably the human participant's early receptive fields). The greater influence of this linear process for the smaller sizes steepens the summation slope beyond fourth root. These steeper slopes are consistent with other studies that tested in the periphery (Manahilov et al., 2001; Meese and Hess, 2007). However, the stimuli in those studies were not as small as those in which we find the steeper slopes in our current study. The noisy energy model can account for this if the design of those studies did not allow participants to restrict their template to the extent of the stimulus. This would defeat one of the two components that combine to give fourth root summation behaviour, resulting in the quadratic summation found in those two studies. For example, in Manahilov et al. (2001) each block was preceded by a suprathreshold example stimulus to reduce uncertainty. The stimuli in that study had a Gaussian spatial envelope, meaning that the "useful" region of the stimulus for detection at contrast threshold would be much smaller than the extent visible in a higher-contrast presentation. This could frustrate the participant's ability to apply an appropriate template.

In the fovea, although some studies have previously reported fourth root summation slopes (Robson and Graham, 1981; Polat and Norcia, 1998; Polat and Tyler, 1999; Meese and Hess, 2007) in most cases the slopes become shallower as the stimulus size increases. This is expected due to the stimuli growing into regions of the visual field which are less sensitive. Compensating for this inhomogeneity using our Witch Hat model allows us to demonstrate that spatial summation in the fovea, the parafovea, and the periphery follows the same fourth root rule.

4.2 Modelling the spatial summation of contrast to threshold

The fourth root summation behaviour here is consistent with earlier studies that found evidence to support the noisy energy model (Meese, 2010; Meese and Summers, 2012; Baldwin and Meese,

2015). However, we acknowledge that the results from this study may also be explained by alternate models that combine spatial filtering with *another* subsequent process that also leads to fourth root behaviour. For example, the signal detection theory-based probability summation models referenced in the **Introduction** (Tyler and Chen, 2000; Meese and Summers, 2012; Kingdom et al., 2015). However, results from our other recent work favours the noisy energy model over a probability summation model (Meese, 2010; Meese and Summers, 2012; Baldwin and Meese, 2015).

In general, the noisy energy model provides a very good account of the results across the different stimulus sizes and visual field locations. It did this with only one free parameter (controlling the global sensitivity) per participant for the foveal and parafoveal results, and with one extra parameter (an offset to account for inaccuracies in the extrapolated Witch Hat model) for the peripheral results. There are two cases, however, where individual participants showed consistent idiosyncratic deviations from the model prediction. The first is for participant SAW (Figs 3C and 4C), where thresholds were systematically lower than predictions for stimuli that were four cycles high and presented in the parafovea. One possible explanation is an individual difference for this participant involving elongated receptive fields in the parafovea (Wilson and Sherman, 1976) with more extensive short-range linear summation within filter elements.

The second deviation from our singular model is found in participant ASB, where thresholds for the smaller stimuli (one cycle high and up to four cycles wide) were approximately 3 dB higher in the periphery than predicted (Figs 3A and 4A). If this result was seen for all of the stimuli with a height of one cycle then a similar "elongated receptive field" explanation could be offered as was made for SAW. Although the data in ASB's case fall *above* the prediction, whereas for SAW they fell *below* it, this could simply be due to the inaccuracy in the Witch Hat extrapolation that we already need account for with our peripheral offset parameter. Unfortunately, such an explanation does not suffice, as the increased thresholds seen for the one cycle high stimuli presented to ASB in the periphery are not found for *widths* greater than four cycles. One possible explanation can be found in models of summation along "skunk tail" stimuli, which suggest there may be additional, more extended, receptive fields that pool along the stripes of a grating stimulus. Relevant to our result is that this second-order filter stage is proposed to differ in its behaviour in the periphery (e.g. becoming phase-sensitive, where it is phase-insensitive in the fovea, see Chen and Tyler, 1999).

An alternate approach to account for ASB's data would be a constraint on the template stage in our model. For example, if the participant was unable to use a template matched to the stimulus area for these small peripheral stimuli. The integration of additional noise in this mismatched template would reduce the signal-to-noise ratio for small stimuli, causing thresholds to lie above the model predictions (as we found). This would not have been compensated for by the Witch Hat, because this region was not mapped by Baldwin et al. (2012). Such a limitation on the template

stage might result from uncertainty about the location of the stimulus in the periphery, or mandatory signal combination over a minimum summation region (Parkes et al., 2001). Further work would be required to determine whether this minimum summation region simply corresponds to the second-order filter stage proposed by Chen and Tyler (1999). Related to this would be the question of how much of the common summation behaviour we find extends to tasks above threshold contrast. The model presented here is relatively simple, having only to predict performance in detecting targets against a blank background. When target stimuli are detected or discriminated at contrasts *above* the detection threshold we expect inhibitory or "gain control" mechanisms to play a substantial role (e.g. Legge and Foley, 1980; Meese and Baker, 2011; Baldwin et al., 2016). Recent work on summation above threshold contrast in collinear ("skunk tail") patterns suggests that these inhibitory signals may be pooled similarly (same summation rule and spatial extent) to the excitatory signals that dominate performance at threshold (Chen et al., 2019, 2023).

5 Additional information

5.1 Acknowledgements

We would like to thank Rob Summers for the Liberator software used to perform this study. This research was supported by an Engineering and Physical Sciences Research Council (EPSRC, UK) grant to TSM and Mark Georgeson (EP/H000038/1), and by a Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant awarded to ASB (RGPIN-2022-04216).

5.2 Author contributions in CREDIT format

ASB: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Software, Visualization, Writing - Original Draft. **TSM**: Conceptualization, Funding Acquisition, Investigation, Methodology, Project Administration, Supervision, Writing - Review & Editing.

5.3 Potential Conflicts of Interest

The authors have no potential conflicts of interest to declare.

References

Baker, D. H. and Meese, T. S. (2011). Contrast integration over area is extensive: A three-stage model of spatial summation. *Journal of Vision*, 11(14):14.

Baker, D. H., Summers, R. J., Baldwin, A. S., and Meese, T. S. (2022). A psychophysical performance-based approach to the quality assessment of image processing algorithms. *PLOS ONE*, 17(5):e0267056.

Baldwin, A. S. (2013a). Chapter 4: The visual field inhomogeneity in contrast sensitivity. In *Pattern Integration in the Normal and Abnormal Human Visual System*, pages 64–92. Aston University, Birmingham, UK.

Baldwin, A. S. (2013b). Section 7.5.4: The stochastic max-across-templates model. In *Pattern Integration in the Normal and Abnormal Human Visual System*, page 140. Aston University, Birmingham, UK.

Baldwin, A. S., Baker, D. H., and Hess, R. F. (2016). What do contrast threshold equivalent noise studies actually measure? Noise vs. nonlinearity in different masking paradigms. *PLOS ONE*, 11(3)(e0150942):1–25.

Baldwin, A. S. and Meese, T. S. (2015). Fourth-root summation of contrast over area: No end in sight when spatially inhomogeneous sensitivity is compensated by a witch's hat. *Journal of Vision*, 15(15)(4):1–12.

Baldwin, A. S., Meese, T. S., and Baker, D. H. (2012). The attenuation surface for contrast sensitivity has the form of a witch's hat within the central visual field. *Journal of Vision*, 12(11)(23):1–17.

Bonneh, Y. and Sagi, D. (1999). Contrast integration across space. *Vision Research*, 39(16):2597–2602.

Chen, C.-C., Chien, C.-H., and Tyler, C. W. (2023). Suprathreshold length summation. *Journal of Vision*, 23(7):17.

Chen, C. C. and Tyler, C. W. (1999). Spatial pattern summation is phase-insensitive in the fovea but not in the periphery. *Spatial Vision*, 12(3):267–285.

Chen, C.-C., Yeh, Y.-H. C., and Tyler, C. W. (2019). Length summation in noise. *Journal of Vision*, 19(9)(11):1–13.

Corso, J. F. (1963). A theoretico-historical review of the threshold concept. *Psychological Bulletin*, 60:356–370.

De Valois, R. L., Albrecht, D. G., and Thorell, L. G. (1982). Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research*, 22(5):545–559.

Field, D. J., Hayes, A., and Hess, R. F. (1993). Contour integration by the human visual system: Evidence for a local "association field". *Vision Research*, 33(2):173–193.

Foley, J. M., Varadharajan, S., Koh, C. C., and Farias, M. C. (2007). Detection of Gabor patterns of different sizes, shapes, phases and eccentricities. *Vision Research*, 47(1):85–107.

Green, D. M. and Swets, J. A. (1966). *Signal Detection Theory and Psychophysics*. John Wiley & Sons Inc, New York.

Gregory, R. L. and Cane, V. (1955). A Statistical Information Theory of Visual Thresholds. *Nature*, 176(4496):1272–1272.

Hilz, R. and Cavonius, C. (1974). Functional organization of the peripheral retina: Sensitivity to periodic stimuli. *Vision Research*, 14(12):1333–1337.

Howell, E. and Hess, R. (1978). The functional area for summation to threshold for sinusoidal gratings. *Vision Research*, 18(4):369–374.

Kingdom, F. A. A., Baldwin, A. S., and Schmidtmann, G. (2015). Modeling probability and additive summation for detection across multiple mechanisms under the assumptions of signal detection theory. *Journal of Vision*, 15(5):1.

Legge, G. E. and Foley, J. M. (1980). Contrast masking in human vision. *Journal of the Optical Society of America*, 70(12):1458.

Manahilov, V. and Simpson, W. (1999). Energy model for contrast detection: Spatiotemporal characteristics of threshold vision. *Biological Cybernetics*, 81(1):61–71.

Manahilov, V., Simpson, W. A., and McCulloch, D. L. (2001). Spatial summation of peripheral Gabor patches. *Journal of the Optical Society of America A*, 18(2):273.

Mayer, M. J. and Tyler, C. W. (1986). Invariance of the slope of the psychometric function with spatial summation. *Journal of the Optical Society of America A*, 3(8):1166.

Meese, T. S. (2010). Spatially extensive summation of contrast energy is revealed by contrast detection of micro-pattern textures. *Journal of Vision*, 10(8):14–14.

Meese, T. S. and Baker, D. H. (2011). Contrast summation across eyes and space is revealed along the entire dipper function by a "Swiss cheese" stimulus. *Journal of Vision*, 11(1):23–23.

Meese, T. S. and Hess, R. F. (2007). Anisotropy for spatial summation of elongated patches of grating: A tale of two tails. *Vision Research*, 47(14):1880–1892.

Meese, T. S. and Summers, R. J. (2007). Area summation in human vision at and above detection threshold. *Proceedings of the Royal Society B: Biological Sciences*, 274(1627):2891–2900.

Meese, T. S. and Summers, R. J. (2012). Theory and data for area summation of contrast with and without uncertainty: Evidence for a noisy energy model. *Journal of Vision*, 12(11):9–9.

Nachmias, J. (1981). On the psychometric function for contrast detection. *Vision Research*, 21(2):215–223.

Parkes, L., Lund, J., Angelucci, A., Solomon, J. A., and Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nature Neuroscience*, 4(7):739–744.

Pointer, J. and Hess, R. (1989). The contrast sensitivity gradient across the human visual field: With emphasis on the low spatial frequency range. *Vision Research*, 29(9):1133–1151.

Polat, U. and Norcia, A. M. (1998). Elongated physiological summation pools in the human visual cortex. *Vision Research*, 38(23):3735–3741.

Polat, U. and Tyler, C. W. (1999). What pattern the eye sees best. *Vision Research*, 39(5):887–895.

Prins, N. and Kingdom, F. A. A. (2018). Applying the Model-Comparison Approach to Test Specific Research Hypotheses in Psychophysical Research Using the Palamedes Toolbox. *Frontiers in Psychology*, 9:1250.

Quick, R. F. (1974). A vector-magnitude model of contrast detection. *Kybernetik*, 16(2):65–67.

Robson, J. and Graham, N. (1981). Probability summation and regional variation in contrast sensitivity across the visual field. *Vision Research*, 21(3):409–418.

Rovamo, J., Mustonen, J., and Näsänen, R. (1994). Modelling contrast sensitivity as a function of retinal illuminance and grating area. *Vision Research*, 34(10):1301–1314.

Sachs, M. B., Nachmias, J., and Robson, J. G. (1971). Spatial-Frequency Channels in Human Vision*. *Journal of the Optical Society of America*, 61(9):1176.

Summers, R. J. and Meese, T. S. (2009). The influence of fixation points on contrast detection and discrimination of patches of grating: Masking and facilitation. *Vision Research*, 49(14):1894–1900.

Swets, J. A. (1961a). Detection theory and psychophysics: A review. *Psychometrika*, 26(1):49–63.

Swets, J. A. (1961b). Is There a Sensory Threshold?: When the effects of the observer's response criterion are isolated, a sensory limitation is not evident. *Science*, 134(3473):168–177.

Tanner, W. P. and Swets, J. A. (1954). A decision-making theory of visual detection. *Psychological Review*, 61(6):401–409.

Tyler, C. W. and Chen, C.-C. (2000). Signal detection theory in the 2AFC paradigm: Attention, channel uncertainty and probability summation. *Vision Research*, 40(22):3121–3144.

Wallis, S. A., Baker, D. H., Meese, T. S., and Georgeson, M. A. (2013). The slope of the psychometric function and non-stationarity of thresholds in spatiotemporal contrast vision. *Vision Research*, 76:1–10.

Wilson, J. R. and Sherman, S. M. (1976). Receptive-field characteristics of neurons in cat striate cortex: Changes with visual field eccentricity. *Journal of Neurophysiology*, 39(3):512–533.