Probability summation for multiple patches of luminance modulation

T.S. Meese a,*, C.B. Williams b

a Vision Sciences, Aston University, Aston Triangle, Birmingham B4 7ET, UK
b Department of Ophthalmology, McGill Vision Research, McGill University, 687 Pine Avenue West [H4-14], Montreal, Quebec, Canada H3A 1A1

Received 28 September 1999; received in revised form 8 February 2000

Abstract

When components of a compound pattern stimulate different visual mechanisms, psychophysical performance typically improves by a small amount consistent with probability summation amongst independent detectors. Here we extend previous summation experiments by (i) plotting full psychometric functions; and (ii) using compound stimuli with components that varied in up to three stimulus dimensions: spatial frequency (1, 4, 5 or 11 c/deg), orientation (0°, ±45°), and position. Stimulus components were isolated circular sine-phase patches of grating centred on up to four corners of an imaginary square surrounding a fixation-point. Combinations of component patches produced compound stimuli made from up to 16 components that differed in various combinations of the three stimulus dimensions. Other than when the spatial frequency was 11 c/deg, results were well described using: (i) probabilistic summation of individual psychometric functions; (ii) the Quick pooling formula; and (iii) the signal detection analysis for 2IFC developed by Tyler and Chen (2000) [Signal detection theory in the 2AFC paradigm: attention, channel uncertainty and probability summation (under review)]. We conclude that in general, nonlinear spatial summation is consistent with probabilistic summation across independent detecting mechanisms that vary in spatial frequency (a range of at least 1–5 c/deg), orientation (a range of 90°) and position (a range of at least 24 cycles at 4 c/deg). In further experiments, results were found to be consistent with probability summation for pairs of orthogonally oriented step-edge stimuli and a matrix of randomly oriented 11 c/deg sine-wave patches. This casts doubt on the generality of a recent suggestion that local interactions between colinearly oriented detectors within a spatial neighbourhood of around four cycles may contribute to nonlinear spatial summation [Bonneh & Sagi, 1998; Vision Research, 38, 3541–3553]. © 2000 Elsevier Science Ltd. All rights reserved.

Keywords: Probability summation; Contrast detection; Spatial vision; Independent detectors

1. Introduction

When the size of a spatially modulated luminance pattern is increased, or its number of (sufficiently distant) Fourier components is increased, contrast sensitivity for the stimulus typically improves. The observed summation is not linear, but approximates the fourth-root of the number of similarly sensitive detectors that have been stimulated and for this reason is sometimes called fourth-root summation. In previous studies (e.g. Quick, 1974; Graham, Robson & Nachmias, 1978; Legge, 1978a,b; Robson & Graham, 1981; Williams & Wilson, 1983), results have been very well described by a model in which the outputs of independent detectors of spatial contrast are combined probabilistically: as more detectors are stimulated, then the probability of detecting the stimulus increases because there is a greater probability that at least one detector will ‘see’ the stimulus. (Sachs, Nachmias & Robson, 1971) referred to this as the ‘inclusive or’ rule). Despite the wide success of this parameter-free model, its theoretical standing has been widely criticized. The first prob-

* Corresponding author. Tel.: + 44-121-3593611; fax: + 44-121-3334220. E-mail address: meese@aston.ac.uk (T.S. Meese)
lem is that Robson and Graham’s analysis assumes a high-threshold model of the detection process which is inconsistent with some experimental data (e.g. Swets, Tanner & Birdsell, 1961; Nachmias, 1981; Gescheider, 1985). This weakness has been addressed by Pelli (1985), who proposed that if the observer were highly uncertain about which detectors were being stimulated, then probability summation could be placed in a signal detection framework and survive the downfall of high threshold theory (HTT). (Also see the comments on Tyler and Chen (2000) in Section 5). Nevertheless, while the theoretical basis of Robson and Graham’s (1981) analysis is thought to be flawed, its established success in fitting experimental data provides a useful starting point for the analysis and discussion in this paper. However, rather than referring to ‘probability summation’ — the theoretical basis of the analysis, we refer instead to ‘HTT summation’ — the summation predicted by high threshold theory and the analysis of Robson and Graham (see Graham, 1989 for details).

The initial motivation for the work presented here was to fill a gap in the experimental literature on nonlinear summation for achromatic spatial patterns. Previous studies have been of two types: Fourier domain and spatial domain. In the first type, compound spatial stimuli have been created by adding together simple sine-wave patterns of different frequencies (e.g. Graham & Nachmias, 1971) or different orientations (e.g. Georgeson & Shackleton, 1994), though compound stimuli have rarely contained more than three stimulus components (Graham, 1980). In the second type, compound stimuli have been created by extending the size of a single patch of grating, either by increasing the number of cycles (e.g. Legge, 1978b; Robson & Graham, 1981) or the length of the stimulus bars (e.g. Howell & Hess, 1978). An alternative spatial domain approach, and one that is perhaps a closer analogue to that in the Fourier domain, is to manipulate the number of spatially discrete patches of luminance modulation. This type of experiment has received little previous attention and may be particularly important in the light of growing evidence for spatially lateral interactions (Polat & Sagi, 1993). Finally, models of spatial vision have been proposed incorporating nonlinear contrast summation across position, spatial frequency and orientation (e.g. Wilson & Bergen, 1979; Watson & Solomon, 1997; Watson, 2000), but experiments have not been performed using stimuli that have been manipulated along all three of these dimensions simultaneously. Thus, to combat the above limitations and extend research on threshold spatial vision, psychometric functions and contrast detection thresholds were measured for (a) individual component patches that varied along the dimensions of position, orientation and spatial frequency, and (b) several different compound stimuli containing up to 16 of the components in (a).

Since we first reported many of the experiments presented here (Meese & Williams, 1998a, b), Bonneh and Sagi (1998) have published a series of experiments assessing the detectability of multi-Gabor element displays. While our own results are broadly consistent with fourth-root summation, Bonneh and Sagi (1998) report that for some spatial configurations, theirs are not. We address this issue further in three supplementary experiments and find little evidence for configurational constraints on nonlinear summation.

2. General methods

Stimuli were stored in the framestore of a VSG2/3 and their presentation was controlled by a Pentium PC. Stimuli were displayed on either an Eizo F553-M monitor with mean luminance of 66 cd/m² at a frame rate of 120 Hz, or a NEC MultiSync XP17 monitor with mean luminance of 69 cd/m² and frame rate of 100 Hz. Look-up tables were used to perform gamma correction of the display monitors and the framestore was operated in pseudo 12-bit mode, allowing a stimulus with Michelson contrast \( \{c = 100[(L_{\text{max}} - L_{\text{min}})/(L_{\text{max}} + L_{\text{min}})]\} \) of 0.5% to be represented by 16 grey levels. An exception to this was Experiment S2 where a frame interleaving technique was used. This improved the grey-level resolution by a factor of two. Component contrast is reported in dB, given by 20 log\((c)\). Thus, 0 dB corresponds with 1%.

A temporal 2IFC technique was used and observers detected the presence of stimulus components and compounds (see below) in randomly interleaved trials (Graham, 1989). Thus, on each trial, observers were unaware of whether the stimulus was a component or a compound and so, presumably, on average they monitored the same detecting mechanisms for all of the stimuli used within an experimental session. In Experiments 1A and S3, a ‘three-down, one-up’ randomly interleaved staircase procedure (Cornsweet, 1962; Wetherill & Levitt, 1965; Meese, 1995) was used to control stimulus contrast, whereas in Experiments 1b, 2, 3, S1 and S2 a method of constant stimuli (MCS) was used. In experiments where thresholds are reported these were estimated using on-line probit analysis (Finney, 1971). For the staircase data, estimates were based on about 100 trials (McKee, Klein & Teller, 1985) and mean thresholds and standard errors were calculated from up to six estimates. For the MCS data, probit analysis was performed on data from 200 trials (160 in Exp S2), and thresholds and standard errors were calculated from up to seven estimates. In all experiments, auditory feedback was used to indicate the correctness of response and the two temporal intervals were marked by short tones simultaneous with the onset of the stimulus.
In several experiments, stimuli consisted of up to four spatially distinct patches of grating or plaid placed at equal distances from a fixation point where each patch was windowed by a raised cosine function with an extended central plateau. All stimulus components were presented in sine-phase with the centre of their patch locations to ensure that none of the stimuli contained a DC component of luminance. The stimulus configurations and dimensions for Experiments 1 and 2 are shown in Figs. 1 and 2, respectively. The dimensions in the figures refer to the boundaries of the half height of the spatial window. In all experiments, stimulus duration was 100 ms. In Experiment 1, stimulus components had a spatial frequency of 1 c/deg and their orientations were vertical. In Experiments 2 and 3, spatial frequency could be either 1 c/deg or 5 c/deg, and orientation could be either 45° or −45° from vertical. In Experiments 1 and 2, no more than a single grating component was placed in each of the four patch locations. In Experiment 3, the stimulus dimensions were the same as those in Experiment 2 (Fig. 2), but up to four components with different orientation and/or spatial frequency were superimposed in up to all four patch locations.

Details of Experiments S1, S2 and S3 are presented later in the paper.

In all experiments other, than Experiment S3, a small dark fixation point was placed at the centre of the display throughout the experiment.

Data were gathered from four observers: the two authors (TSM & CBW), another psychophysically experienced observer (RFH) and a naive observer with limited psychophysical experience (CHD). All observers had normal or optically corrected to normal vision. TSM and CHD carried out most of their experiments using a chin and forehead rest at viewing distances of either 83 or 228 cm. CBW and RFH (and TSM in Experiment 1a) sat at a viewing distance of 114 cm.

3. Results

3.1. Experiment 1a

In Experiment 1a, seven different pairs of interleaved staircases tracked the detection thresholds for: (i) each of the four individual patches of grating; (ii) two compounds consisting of pairs of patches placed in opposite corners of the display (i.e. upper right and lower left, and upper left and lower right); and (iii) a compound consisting of all four patches. Results are averaged from up to six experimental sessions and shown in Fig. 3 for TSM and Fig. 4 for RFH. In both figures, the solid line is a prediction for HTT summation derived from the Quick (1974) pooling formula, which on a log–log plot has a slope of $-1/\beta$ (Robson & Graham, 1981), where $\beta$ was estimated from the slopes of the psychometric functions for each observer (see forward to Experiment 3 for details of this estimate). The vertical position of the fit was determined by a least-squares
Fig. 3. Results from Experiment 1a for TSM. Detection thresholds (contrast in dB = 20 log(c)) where c is Michelson contrast in percent), as a function of the number of grating patches in the stimulus. The stimulus configuration was that shown in Fig. 1. The solid line is a prediction for HTT summation. Its slope on a log–log plot is given by \( \frac{1}{\beta} \), where \( \beta \) was estimated from the average slope of the psychometric functions measured for this observer. The vertical offset of the slope was determined by a least-squares routine. Error bars show ±1 SE.

procedure (Press, Flannery, Teukolsky & Vetterling, 1989). In both cases, the model provides an excellent fit to the data. Note that for RFH, compared with TSM, the slope of the psychometric function is shallower (lower value of \( \beta \)), and the amount of both the predicted and measured summation is greater. In a second fit of the model to the threshold data (not shown), \( \beta \) was a free parameter. The fits produced \( \beta = 4.34 \) for TSM and \( \beta = 3.05 \) for RFH. These values are in excellent agreement with the values estimated from the psychometric functions (see Figs. 3 and 4).

3.2. Experiment 1b

In Experiment 1b a method of constant stimuli was used in order that full psychometric functions could be derived for each stimulus (Graham, 1980). In a single experimental session, sensitivity was measured for individual grating patches in each of the four different locations as well as a compound stimulus made from the simultaneous presentation of all four patches of grating (Fig. 1). Psychometric functions are shown in Fig. 5 for TSM and in Fig. 6 for CBW (in all figures plotting psychometric functions, error bars are shown for the compound stimuli only and represent ±1 SE where this is larger than the data symbol). Data are from four or five sessions and each session consisted of 1000 trials (200 trials for each of the four components and 200 trials for the 4-patch compound). Data were corrected for guessing using Abbott’s formula:

\[
P_d = \frac{P_c - g}{1 - g}
\]

(1)

where \( P_d \) is the probability of detecting the stimulus and \( P_c \) is the proportion of correct experimental trials.

Fig. 5. Psychometric functions and predictions for Experiment 1b for observer TSM. The stimulus configuration was that shown in Fig. 1. The open symbols are results for each of the four individual patches of grating and the filled symbols are for the compound stimulus containing all four patches. Error bars show ±1 SE and for clarity are shown for the four patch condition only. The dashed and thick solid curves are predictions for HTT summation (see text for details). Component contrast is in dB and given by 20 log(c), where c is Michelson contrast in percent. Thus, 0 dB corresponds with 1%.
Variable $g$ is the guess rate which for 2IFC is 0.5. If performance is worse than the guess rate then $P_d$ becomes negative, which is problematic for the analysis. In such circumstances, this problem was avoided by resetting $P_d$ to zero (see below for further discussion):

If $P_d < 0$ then $P_d = 0$  

(2)

In Fig. 6, the data for CBW have also been corrected for ‘finger errors’ (pressing the wrong key by mistake). This was done because even at the highest contrast level used for this observer (3 dB (1.4%)), performance was less than optimum (detection rate = 80% after correcting for guessing) even though in identical but more slowly paced control sessions (Meese & Williams, 1998a) performance reached 100% for the same contrast level. Thus, at all contrast levels, the experimental data for CBW were divided by 0.8 to compensate for finger errors and produce the results shown in Fig. 6.

In Figs. 5 and 6, the thick curves show predictions for HTT summation. For completeness, we state the assumptions behind the theoretical basis for this analysis. These are: a high threshold model of detection, negligible false positive responses in 2IFC, independent detection of each patch and probability summation between detectors (Graham, 1989). From standard analysis of probabilities, the probability of detecting at least one of the patches in the 4-patch compound is given by

$$P_{1,2,3,4} = (1 - Q_1 \cdot Q_2 \cdot Q_3 \cdot Q_4),$$

where $Q_1$, $Q_2$, $Q_3$ and $Q_4$ are the probabilities of not detecting each of the four patches. Substituting $1 - P_i$ for $Q_i$ and expanding gives

$$P_{1,2,3,4} = (P_1 + P_2 + P_3 + P_4)$$

$$- (P_1P_2 + P_1P_3 + P_1P_4 + P_2P_3 + P_2P_4 + P_3P_4)$$

$$+ (P_1P_2P_3 + P_1P_2P_4 + P_1P_3P_4 + P_2P_3P_4)$$

$$- P_1P_2P_3P_4$$

(3)

where $P_1$, $P_2$, $P_3$ and $P_4$ are the probabilities of detecting each of the four individual patches and can be estimated from experimental results.

The data were handled in two different ways. For the dashed curves, data from different experimental sessions were collapsed before applying Eq. (3). For the thick solid curves, Eq. (3) was applied to the data from each individual session and then the average prediction was calculated. The nonlinearity of Eq. (3) means that these two treatments can produce slightly different predictions. In Figs. 5 and 6, analysis is shown also for situations where only the three and two most detectable stimulus components contributed to the analysis.

For both observers, at least three of the most detectable components were required in the analysis to provide a good fit to the 4-patch compound condition (filled circles). For CBW, the slight overestimation by the thick solid curve at low contrast may be due to the use of Eq. (2), which has the overall effect of slightly overestimating average performance when the probability of seeing the stimulus is very low. The average predictions (dashed curves) suffer much less from this because averaging a large amount of data makes it less likely that $P_d$ is less than zero. However, this averaging process does have the disadvantage that any fluctuations in threshold location between experimental sessions will result in an artificial shallowing of the psychometric function which will result in an overestimation in HTT summation. Fortunately, for fluctuations that are likely to be encountered in many psychophysical experiments, this source of error is negligible (Meese & Williams, 1998b; Meese 2000b).

3.3. Experiment 2

In Experiment 2 the stimulus was modified to reduce the possibility of linear summation of two or more
components within a single detector. This was achieved in the spatial domain by reducing the size of the test patches and in the Fourier domain by using different combinations of orientation and spatial frequency for each patch (see Fig. 2).

To equate approximately the detectability of the four stimulus patches, the contrasts of the 5 c/deg patches were set 6 dB (factor of 2) higher than the 1 c/deg patches for TSM. For CHD, based on preliminary measurements, the contrast of each patch was adjusted in an attempt to make them equally detectable.

Results are shown for TSM in Fig. 7 and a naive observer, CHD, in Fig. 8. The predictions were calculated in the same way as in Experiment 1b. Once more, the fit is fairly good for both observers when more than the two most detectable components are considered, though for CHD, the data are also fairly well described by HTT summation for only the two most detectable components.

3.4. Experiment 3

In Experiments 1 and 2, compound stimuli contained up to four component patches in different spatial locations and the results were broadly consistent with HTT summation. In Experiment 3 the number of components in the compound was extended to as many as 16 to investigate whether HTT summation occurs for stimulus components in different locations, different spatial frequencies and different orientations. Thus, detection thresholds were measured for 16 different components (4 positions × 2 orientations × 2 spatial frequencies), and four compounds made from 16, 8, 8 and 4 of these components (see Fig. 9). Like in Experiment 2, the contrast of the 5 c/deg components were set 6 dB higher than those of the 1 c/deg components in an attempt to equate detectability. Because of the large number of components used in this experiment it was not practical to interleave all of the experimental conditions within each experimental session. Instead, a compromise was made where, in each session, trials were interleaved for one of the four compound stimuli plus four of the 16 component stimuli, and the order of sessions was performed in randomised blocks. In all sessions, component stimuli occupied each of the four patch locations, had each of the two orientations and each of the two spatial frequencies. The compound stimuli consisted of: (1) four components of both orientations and both spatial frequencies in the upper left spatial location; (2) eight components at all locations and both spatial frequencies, but only the left oblique orientation; (3) eight components at all locations and both orientations, but only at 1 c/deg; and (4) all 16 components. Thresholds for each of the components and the four compounds are shown in Fig. 9.

It is of passing interest that although sensitivity is fairly uniform across orientation and position at 1 c/deg, at 5 c/deg the four components to which sensitivity was highest were those with orientations radiating out from the fixation point. This is generally consistent with the ‘meridional resolution effect’ reported by Rovamo, Virso, Laurinen and Hyvarinen (1982).

3.5. Quick pooling formula

It is impractical to use an extended version of Eq. (3) to make predictions for HTT summation when the number of stimulus components is large. If only thresholds are considered, however, then the Quick (1974) pooling formula can be used. This is given by:

\[ S_c = \left( \sum S_i^{\beta} \right)^{1/\beta}, \]  

where \( S_c \) is the sensitivity to the compound stimulus and \( S_\text{i} \) is the sensitivity to its \( i \)th component alone. Typically, the parameter \( \beta \) is estimated from the slope of the psychometric function fit by a Weibull function. An alternative function, and the one used in this study, is the log-normal ogive which was fit to the detection data using on-line probit analysis. The spread parame-

---

\[^2\text{This formula makes the same predictions as that of Eq. (3) if the psychometric function is a Weibull function (Quick, 1974). However, other sigmoidal functions deviate only slightly from the Weibull function, and when the number of components in the analysis is small the difference in predictions is small. Furthermore, a reanalysis of Experiments 1b and 2 using the Quick pooling formula (Meese & Williams, 1998b) produced a similar outcome to the analysis presented here. See Wilson and Bergen (1979), Robson and Graham (1981), Tyler and Chen (2000), Meese (2000b) and Section 5 of this paper for further comments and discussion.}\]
4. Supplementary experiments

In a recently published paper, Bonneh and Sagi (1998) used stimuli made from small patches of high spatial frequency grating and reported nonlinear summation of a magnitude consistent with HTT summation only when the patches were colinear and closely spaced. For other stimulus configurations, the magnitude of summation was found to be less. This led the authors to suggest that the observed summation may be due to facilitatory interactions between nearby colinearly oriented detectors. The results from the following three supplementary experiments show that the constraints reported by Bonneh and Sagi (1998) are not general.

4.1. Experiment S1

To further examine the issue of proximity we repeated Experiment 1b but used the much smaller Gabor stimulus patches shown in Fig. 10. Each patch is a sine-wave grating multiplied by a Gaussian window function with a full-width at half-height of 1.42 cycles (similar to Bonneh & Sagi’s parameter of 1.39 cycles). The sine-waves had a spatial period of ten pixels per cycle. The horizontal and vertical distances between the centres of the patches is 24 cycles. This is considerably greater than the distance of around three cycles beyond which Bonneh and Sagi (1998) found a reduction in nonlinear summation. It is also greater than the distance of 12 cycles, at which a facilitatory effect of flanking mask patches on a central test patch was found to be abolished (Polat & Sagi, 1993). In one condition, the spatial frequency was 4 c/deg, the full width at half height of our Gaussian envelope was 0.36° and the vertical spacing between patch centres was 6°. Based on preliminary observations, the contrasts of the two 4 c/deg patches in the lower visual field were set 1.5 dB lower than the other two patches in an attempt to equate detectability. In a second condition, contrasts were equal and the viewing distance was increased from 83 to 228 cm, to give a spatial frequency of 11 c/deg.

At 4 c/deg (Fig. 11) the pattern of results was very much the same as in Experiment 1b: the results for the 4-patch compound stimulus were well described by HTT summation for the four stimulus components (coarse dashed curve), fewer components in the analysis (fine dashed curve) tending to underestimate summation. This result is not consistent with Bonneh and Sagi’s (1998) suggestion that nonlinear summation of multiple components may be due to local interactions because the distance between the patches is consider-
ably greater than the distance over which interactions are thought to take place (e.g. Polat & Sagi, 1993).

At 11 c/deg (Fig. 12), summation was a little less than HTT summation. Note, however, that in our own experiment, two of the components were barely detectable, even at a contrast of 39 dB (89%), meaning that the differences in predictions for two, three and four components is small (see dashed curves in Fig. 12). The apparent contrast of these components seemed higher at the beginning of each experimental session, suggesting that the low sensitivity may be due to local adaptation or difficulties in maintaining accommodation. Another possibility is that stimulus uncertainty might increase during the experiment, leading to high thresholds (Pelli, 1985).

To avoid the problem of low sensitivity to the 11 c/deg grating patches, the experiment was repeated with the same stimulus components, but with patches spaced only 12 cycles apart so as to stimulate a more sensitive region of the retina (Robson & Graham, 1981). The results are shown in Fig. 13. Once again, a small amount of summation is evident but it is less than that predicted by HTT summation (coarse dashed curve) for four stimulus components.

4.2. Experiment S2

In Experiment S1 we addressed the issue of proximity raised by Bonneh and Sagi (1998). In the next experiment we address the issue of colinearity. We measured detection thresholds for each of a pair of orthogonally oriented step-edge stimuli and their compound shown in Fig. 14 (see figure legend for stimulus dimensions), to examine the possibility that nonlinear spatial contrast summation is consistent with HTT summation only for colinearly oriented contours.

Results are shown in Fig. 15. The HTT summation prediction (coarse dashed curve) using Eq. (3) (with $P_3$ and $P_4$ set to zero), provides an excellent account of the data. This result is in general agreement with that of Experiment 2: luminance modulations do not need to be at similar orientations for HTT summation to occur.

We repeated this experiment for the same observer, replacing the vertical stimulus patch with another horizontal stimulus patch. Results were similar to those in Fig. 15 (for example, with $\beta = 3.64$ estimated from the psychometric functions, Eq. (4) underestimated summa-
tion by less than 0.2 dB), again suggesting that there is nothing special about colinearity and nonlinear summation.

4.3. Experiment S3

Bonneh and Sagi (1998) presented the results of several contrast detection experiments which led them to suggest that nonlinear summation was constrained by stimulus configuration. While the above results indicate that this is not generally so for proximity (Experiment S1) nor colinearity (Experiment S2), we felt further examination of another one of their experiments was worthwhile. Bonneh and Sagi (1998) pointed out that if nonlinear summation were due to HTT summation, then randomising the orientations of the component patches should have no effect on summation. They found, however, much more summation amongst 576 co-oriented Gabor patches than amongst the same number of randomly oriented patches. Prompted by this finding we performed a similar experiment using 1, 49 and 81 stimulus patches, which were similar to those used in Experiment S1, and had a spatial frequency of 11 c/deg. The stimulus configurations for the 49-patch stimuli are shown in Fig. 16. In Fig. 16a, the patches are co-oriented and all have the same phase. In Fig. 16c, the configuration is the same as that in Fig. 16a, but the phase of alternate patches is shifted by 180°. In Fig. 16b, the orientations are randomly selected from the following eight orientations: 90° (horizontal), −67.5°, −45°, −22.5°, 0° (vertical), 22.5°, 45° and 67.5°. The randomisation of patch orientation was performed independently on each experimental trial. The 81-patch stimuli were similar to those in Fig. 16a,b but contained nine rows and columns of patches. For all of these stimulus configurations, the central patch was always vertical and in sine-phase. Data were gathered using interleaved staircases in three series of experimental sessions. In the first series, thresholds were measured for the three 49-patch configurations (Fig. 16a,b,c) and a single vertical patch. In the third series, thresholds were measured for a single oblique patch and a single vertical patch. In all cases, the fixation point was removed 300 ms before the onset of the stimulus.

Detection thresholds (± 1 SE) are shown in Fig. 17 relative to that of the single vertical patch measured within the common series of sessions. In each of the multiple-patch conditions the detection threshold is between 4 and 6 dB less than for the single central patch, representing substantial summation of some kind. A detailed analysis of this experiment is difficult because we did not measure sensitivity to each patch orientation in each of the 81 different locations, but we suspect that in both this experiment and that of Bonneh and Sagi (1998), the rapid decrease in sensitivity with increasing distance from the fovea at high spatial frequencies (e.g. Robson & Graham, 1981; also see Experiment S1) would result in few of the stimulus patches contributing...
to detection. The similarity in thresholds for the 49-patch conditions and the 81-patch conditions are consistent with this view. The horizontal dashed lines in Fig. 17 show HTT summation for nine (3 rows x 3 columns) and 18 patches for which sensitivity is equal, using Eq. (4) and \( \beta = 4.46 \) estimated from the psychometric functions measured in the present experiment. These predictions span the range of observed summation fairly well. However, for both 81-patch and 49-patch conditions there is about 1.5 dB more summation when the stimulus patches were all vertical than when they were randomly oriented (though note the large standard error bars). The direction of this result is the same as for Bonneh and Sagi (1998), though we note that for Bonneh and Sagi, the difference (averaged across three observers) between their co-oriented and random conditions is closer to 3 dB. One possibility is that the summation is partly due to linear summation within filters with large oriented receptive fields (e.g. Polat & Tyler, 1999) which would not be sensitive to the randomised patches. However, when the phase of alternate vertical patches was shifted by 180°, sensitivity was unchanged (Fig. 17), suggesting that large vertically oriented linear filters were not being used. On the other hand, nonlinear phase-insensitive mechanisms with large oriented receptive fields could be responsible for some of the summation seen here and by Bonneh and Sagi (1998). This result and conclusion is consistent with that of Chen and Tyler (1999), who found sensitivities were alike for strips of alternating-phase and in-phase patches of grating in the fovea. A different possibility follows from our finding that sensitivity was about 1.5 dB less for a single obliquely oriented patch than for a single vertical patch (Fig. 17): the well known oblique effect (Campbell, Kulikowski & Levinson, 1966). It is perhaps not surprising that sensitivity is a little less for the random pattern than it is for the vertical pattern given that the random pattern contains patch orientations to which the visual system is less sensitive. However, if we assume that for the random pattern, the effective contrast of half of the components was 1.5 dB less than for the vertical pattern, then this effect would account for a difference between the random and vertical patterns of only 0.6 dB (using Eq. (4)). Furthermore, the control manipulations performed by Bonneh and Sagi (1998) make this account seem less likely in their case. It remains unclear why more summation is found for co-oriented stimulus patches than for randomly oriented patches, though pooling by second-order oriented mechanisms that are insensitive to phase is one possibility. Nevertheless, our results with randomly oriented and co-oriented gratings are consistent with HTT summation (or similar) amongst a subset of between 9 and 18 (groups of) detectors.

5. Discussion

In this paper we have presented the results from a series of experiments investigating the detectability of multi-component stimuli. First, we discuss our results within the context of HTT summation and consider their meaning for the lateral interaction hypothesis of Bonneh and Sagi (1998). Then, we reconsider our results in the light of the 2IFC signal detection analysis of Tyler and Chen (2000), also see Tyler (1997) and Tyler (2000).

5.1. HTT summation and lateral interactions

The majority of the results presented in this paper are consistent with HTT summation. That is, the detectability of a compound stimulus can be predicted from the detectability of the component stimuli with no free parameters using Eqs. (3) and (4). This suggests that mechanisms selective for different positions, orientations and spatial frequencies can contribute simultaneously to stimulus detection. Our data are not generally consistent with Bonneh and Sagi’s (1998) finding that HTT summation occurs only for spatially adjacent, or nearby detectors (e.g. Fig. 11) or those with colinearly oriented receptive fields (e.g. Fig. 15). This makes a general explanation of nonlinear summation in terms of selective lateral interactions (Usher et al., 1999) seem unlikely. For example, in Usher et al.’s model, lateral interactions exist primarily between co-oriented detectors within a spatial neighbourhood of around nine cycles (see Usher et al., 1999 for details). These interactions result in approximately fourth-root summation for appropriately constrained stimulus configurations, but, in contradiction to our results, would produce little or no summation for the stimulus configurations in Experiments S1 and S2.

Our only experiment in which summation was clearly different from HTT summation was when the test spatial frequency was 11 c/deg (though perhaps also naive subject CHD in Experiment 2). In this case, observed summation was less than HTT summation. This occurred for a spatial frequency similar to that used by Bonneh and Sagi (12.5 c/deg), where summation was found to depend upon spatial configuration. Thus, one way in which some of Bonneh and Sagi’s (1998) results might be reconciled with ours, is to suppose that the configurational effects are specific to high spatial frequencies. But how could this be? There are at least three distinct possibilities. First, if the constraints in Usher et al.’s (1999) model were relaxed at moderate and low spatial frequencies (e.g. if interactions were to take place between all mechanisms within low and moderate spatial frequency bands), then their model would be consistent with much of the data presented here. It is unclear, however, what the purpose
of such unspecific interactions would be. A second possibility is that probability summation occurs at low and mid spatial frequencies, but is less prevalent at high spatial frequencies (e.g. filter noise becomes correlated at higher spatial frequencies). Configuration specific effects are then explained by supposing interactions similar to those proposed by Usher et al at high spatial frequencies only. A third and very different possibility supposes no interactions at all, but that observers are more uncertain about the stimulus when small (widely spaced) stimulus patches are used in unusual spatial configurations. This would steepen the slope of the instantaneous psychometric function and decrease summation (Pelli, 1985; Tyler & Chen, 2000), though the required increase in slope might need considerable change in uncertainty. At first sight, this proposal seems inconsistent with the data because the psychometric functions at high spatial frequencies (see Figs. 12 and 13) are not unusually steep. However, it is plausible that these psychometric functions are made artificially shallow by sources of variability that are correlated between detectors. The variability in detector output produced by small eye-movements and fluctuations in accommodation are two plausible examples that would manifest themselves most severely at high spatial frequency.

5.2. Signal detection theory and summation

Tyler and Chen (2000) have derived summation predictions for the 2IFC paradigm in a signal detection framework. In their work, a linear transducer is assumed for simplicity, and predictions and approximations are shown for a variety of other assumptions including additive noise. One crucial feature of the analysis is that to achieve a steep psychometric function, like those observed in the experiments, it is necessary that uncertainty is high. Uncertainty depends upon the ratio $m:n$ (Pelli, 1985), where $m$ is the number of detecting mechanisms that are monitored, and $n$ is the number of equally sensitive detecting mechanisms that are stimulated. Tyler and Chen show that when $m$ is large (at the time of going to press, in their Fig. 10, $m = 1000$) and when the size of the attention window is fixed, thresholds decrease with a slope of approximately $-1/4$ (on a log–log plot) over a range of $n$ from 1 to 100. Thus, as illustrated in the next section, the 2IFC results reported here are broadly consistent with signal detection theory assuming uncertainty and additive noise.

5.3. Fourth-root summation

Several authors (e.g. Laming, 1986; Graham, 1989; Bonneh & Sagi, 1998) have suggested an alternative interpretation of Eq. (4); namely, nonlinear summation between the outputs of independent detectors, where $\beta$ determines the exponent of summation and has a (typical) value of 4. In this case, Eq. (4) can be treated as a descriptive nonlinear summation model of the data and no assumptions about the shape or slope of the psychometric function are necessary. Furthermore, this interpretation is also consistent with more specific ideas about probability summation and the assumptions of signal detection theory as outlined above. Note that in either case, the success of Eq. (3) is coincidental. Comparisons are made in Fig. 18, where the compound data from Experiments 1b, 2, S1 and S2 are replotted with predictions using Eq. (3) (coarse dashed curves) and Eq. (4) with $\beta = 4$ (thin solid curves). The differences in predictions are slight, though in detail, the nonlinear summation model (Eq. (4)) fares a little better in the high frequency conditions (bottom panels), while the HTT summation predictions (Eq. (3)) fare a little better with several of the other conditions (e.g. the step-edge stimuli). The similarity of the fits is presumably due to the similarity of the empirical psychometric functions to a Weibull function with slope parameter of 4.

On the other hand, one particularly striking finding in Experiment 1a is the match between the different slopes of the psychometric functions and the different summation slopes for the two observers. This finding is consistent with the analysis of Robson and Graham (1981), but it remains unclear to us whether it can be accommodated by the framework of Tyler and Chen. Certainly, this result, and others (e.g. Williams & Wilson, 1983), where the summation slope is different from 1/4, is not consistent with a strict fourth-root summation interpretation of Eq. (4).

6. Summary

The data presented in this paper are broadly consistent with probability summation amongst independent contrast detectors selective for different spatial frequencies, orientations and position. At moderate and low spatial frequencies (i.e. $\leq 5$ c/deg), summation is not constrained by stimulus configuration in terms of either proximity or contour alignment (Bonneh & Sagi, 1998). These conclusions are consistent with the 2IFC signal detection and probability summation analysis of Tyler and Chen (2000) as well as a more general fourth-root summation model. However, unlike the Robson and

---

3To implement this model it was necessary to interpolate and extrapolate between and beyond data points to estimate the stimulus contrast for any particular probability of detecting each of the stimulus components. At the low contrast end of the scale, where extrapolation is unsafe, this estimate was arbitrarily set to the lowest (normalised) stimulus contrast used. This was done for all probabilities of component detection that were lower than for those cases where interpolation could be performed.
Graham analysis for HTT summation, the fourth-root summation model is unable to account for the interrelation between summation and the slope of the psychometric function found in Experiment 1a.

Acknowledgements

This work was supported in part by EPSRC of UK (GR/M10854), NSERC of Canada (OGPOO 46528) and MRC of Canada (mt 10818). Several of the experiments in this paper were reported at the 1997, 1998 and 1999 AVA Christmas Meetings at Aston, UK (Meese & Williams, 1998a; Meese, 2000a) and ECVP 1998 at Oxford UK (Meese & Williams, 1998b). We would like to thank Christopher Tyler and an anonymous reviewer for helpful exchanges and comments on the ms.

References


