Pattern Integration in the Normal and Abnormal Human Visual System

Alexander Scott Baldwin

Doctor of Philosophy

Aston University

March, 2013

Thesis Summary

Aston University

Pattern Integration in the Normal and Abnormal Human Visual System

Alexander Scott Baldwin

PhD candidate in Neurosciences

Submitted: 2013

The processing conducted by the visual system requires the combination of signals that are detected at different locations in the visual field. The processes by which these signals are combined are explored here using psychophysical experiments and computer modelling. Most of the work presented in this thesis is concerned with the summation of contrast over space at detection threshold. Previous investigations of this sort have been confounded by the inhomogeneity in contrast sensitivity across the visual field. Experiments performed in this thesis find that the decline in log contrast sensitivity with eccentricity is bilinear, with an initial steep fall-off followed by a shallower decline. This decline is scale-invariant for spatial frequencies of 0.7 to 4 c/deg. A detailed map of the inhomogeneity is developed, and applied to area summation experiments both by incorporating it into models of the visual system and by using it to compensate stimuli in order to factor out the effects of the inhomogeneity.

The results of these area summation experiments show that the summation of contrast over area is spatially extensive (occurring over 33 stimulus carrier cycles), and that summation behaviour is the same in the fovea, parafovea, and periphery. Summation occurs according to a fourth-root summation rule, consistent with a "noisy energy" model. This work is extended to investigate the visual deficit in amblyopia, finding that area summation is normal in amblyopic observers. Finally, the methods used to study the summation of threshold contrast over area are adapted to investigate the integration of coherent orientation signals in a texture. The results of this study are described by a two-stage model, with a mandatory local combination stage followed by flexible global pooling of these local outputs. In each study, the results suggest a more extensive combination of signals in vision than has been previously understood.

Keywords: *psychophysics*, *spatial vision*, *pattern vision*, *contrast*, *area summation*

Acknowledgements

I primarily wish to acknowledge the assistance and support I have received from my supervisor, Tim Meese. I would also like to thank the other members of the Sensory and Perceptual Systems research group at Aston University. I am particularly grateful to Mark Georgeson, Daniel Baker, Robert Summers and Stuart Wallis for their advice. Daniel Baker assisted with the experiment design for Chapters 4 and 6 of this thesis.

Conducting the research for this thesis involved visiting McGill Vision Research on two occasions (summers of 2011 and 2012). I would like to thank Robert Hess, Kathy Mullen, Jesse Husk, Pi-Chun Huang, and Simon Clavagnier for their hospitality and assistance. Robert Hess and Kathy Mullen provided lab space and equipment that I used to perform my experiments. Jesse Husk, Pi-Chun Huang, and Simon Clavagnier helped me recruit amblyopic observers for Chapter 8 of this thesis. Robert Hess and Jesse Husk assisted with the experiment design and observer recruitment for Chapter 9.

I am also grateful to all of the observers who participated in my experiments, and to my friends and family for their support over the past four years.

Publications

Some of the work published in this dissertation has previously been presented in poster form at scientific conferences:

AVA Spring Meeting 2010 (Liverpool Hope University). Baldwin, A. S., Meese, T. S., and Baker, D. H. (2010). Loss of contrast sensitivity at 4 c/deg depends on eccentricity and meridian but not grating orientation for the central 9 deg of the visual field. *Perception* 39(8):1151.

AVA Christmas Meeting 2010 (Descartes University). Baldwin, A. S., Meese, T. S., and Baker, D. H. (2011). Retinal inhomogeneity and the witch's hat: contrast sensitivity declines as a bi-linear function of eccentricity in each direction. *Perception* 40(1):112.

AVA Spring Meeting 2011 (Cardiff University). Baldwin, A. S., Meese, T. S., and Baker, D. H. (2012). Extensive physiological summation of contrast signals over area revealed by witch's hat compensation for retinal inhomogeneity. *Perception* 41(3):366.

ECVP 2012 (Alghero, Sardinia). Baldwin, A. S., Meese, T. S., and Baker, D. H. (2012). A reevaluation of area summation of contrast with compensation for retinal inhomogeneity. *Perception* 41 (ECVP Abstract Supplement):223.

AVA Christmas Meeting 2012 (UCL). Baldwin, A. S., Husk, J. S., Meese, T. S., and Hess, R. F. (2012). Pooling strategies for the integration of orientation signals depend on their spatial configuration. *Perception* 41 p. 1512.

Additionally, the study in Chapter 4 has been published as a paper:

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 pp. 1-17.

Contents

List of Figures

List of Tables

CHAPTER 1

Introduction

> "We are so familiar with seeing, that it takes a leap of imagination to realize that there are problems to be solved. But consider it. We are given tiny distorted upside-down images in the eyes, and we see separate solid objects in surrounding space. From the patterns of stimulation on the retina, we perceive the world of objects and this is nothing short of a miracle."
> – *Richard Gregory (1966)*

THE human visual system has evolved to perform the difficult task of extracting and interpreting information from the outside world in order to inform our behaviour. The retinal circuitry transduces the patterns of light projected onto it by the eye into signals which are carried and processed by neurones. At the early stages of the visual system neurones are tuned to respond to specific stimulus properties at the retinal location which that neurone corresponds to. The response from the individual neurones are ambiguous however, as although their responses are tuned there are still many possible stimuli that could activate any single neurone. For example in primary visual cortex, a particular neurone will respond to a bar or an edge in an image at a particular orientation and spatial scale. The responses of many of these neurones can be combined to create higher-level detectors which are specific to more complex stimuli such as regular textures or contours. It is this combination of signals over space that is investigated here.

The studies conducted in this thesis investigate the ways in which visual stimuli are processed by using the complementary techniques of psychophysics and computer modelling. The brain can be considered to be a "black box" machine which performs a set of unknown functions on a known set of inputs to produce a measurable set of outputs. The inputs are the stimuli used in experiments and the outputs are the responses of human observers, usually expressed as probabilities of responding in a certain way to a particular stimulus (e.g. detecting a faint pattern). The functions that the brain performs on the visual input can be inferred by the properties of the behavioural output. The range of possible mechanisms that might be used to implement these functions can be guided by what is known from anatomical and physiological

investigations. Developing mathematical and computer models of how this processing might occur allows for predictions to be generated that can be compared against the data. If human observers show similar behaviour to that predicted by a particular computer model, then the processing implemented in that model may be the same as that implemented in the brain.

1.1 This thesis

This thesis details my work in applying the techniques of psychophysics and computer modelling in order to explain how signals from different locations in the visual field are combined into percepts of objects and surfaces extended in space. After this introduction (Chapter 1), the following two chapters are a review of the relevant literature (Chapter 2) and an introduction to the methods that are used in the experiments presented here (Chapter 3). A large part of this thesis is concerned with how the detection of low contrast (faint) signals is aided by adding together samples from different locations in the visual field. One key component in models of these area summation tasks is a map of how the sensitivity to these signals varies across the visual field. In the first experimental chapter (Chapter 4) the purpose was to create a detailed model of this inhomogeneity in contrast sensitivity. This is incorporated into computer models of how threshold contrast should decrease for a stimulus as it increases in size. These models are described in the general modelling chapter (Chapter 5) and are then applied to the empirical results in the subsequent experimental chapters.

In the first of the area summation studies shown here (Chapter 6), the models described in Chapter 5 are tested against data collected using two different stimuli (gratings and "Swiss cheeses"). Conditions are also tested where the stimuli are compensated to counteract the visual field inhomogeneity in contrast sensitivity (based on the work in Chapter 4), with the intention of finding out how large an area observers are able to combine contrast signals over. In Chapter 7 these same techniques are applied to rectangular strips of grating presented in the fovea, parafovea, and periphery, in order to determine whether area summation varies between these different locations. In the next chapter (Chapter 8), area summation is investigated in subjects with amblyopia, using "Battenberg" stimuli. Amblyopia is an acquired disorder of the visual system that is caused by impaired vision in an eye during childhood. In subjects with amblyopia the development of the visual cortex is disrupted. The work presented here investigates whether this results in abnormal summation behaviour. The final experimental chapter (Chapter 9) adapts these area summation methods to provide an investigation of the combination of orientation information across fields of oriented elements. Finally, there is the discussion (Chapter 10) and the appendices (Appendix A-C).

All experimental chapters are prefaced by the "motivation and summary" for that chapter, as well as a brief review of the relevant literature. Redundancy is avoided where possible through the use of references back to the Literature Review, General Methods, and General Modelling, or through references to other experimental chapters.

CHAPTER 2

Literature Review

2.1 Introduction

As one of the fundamental tasks that must be carried out by the visual system, the combination of signals from separate locations in the visual field has been studied extensively in the past. In this chapter I survey the previous literature concerning the summation of contrast over area and set out the general background and motivation for the series of experiments presented in this thesis. Additional summaries are provided of signal combination findings in other domains.

2.2 Basic architecture of early vision

2.2.1 The retina

To a first approximation, the human visual system can be considered to consist of a series of feedforward stages (the visual system does also contain extensive feedback connections, however these are omitted for the sake of brevity here). The processing of visual information from the outside world begins in the eye. Individual photoreceptors in the retina detect light of different intensities and wavelengths incident upon them. These signals are combined through neuronal convergence onto ganglion cells. The geometry and polarity of the connections between the photoreceptors and the retinal ganglion cells (via the intermediate horizontal, bipolar and amacrine cells) bring about their characteristic antagonistic centre-surround receptive fields (Kuffler, 1953). These cells respond to local variations in luminance between the centre and the concentric annulus of their receptive field (either favouring light spots as “on-centre” retinal ganglion cells or dark spots as “off-centre” retinal ganglion cells).

One interpretation of the ganglion cell receptive field structure is that they perform a filtering operation on the retinal image, with the size of the excitatory and inhibitory regions determining their spatial filtering properties. Those with smaller receptive fields respond to finer details (higher spatial frequency), whereas those with larger receptive fields respond to coarser details (lower spatial frequency). The firing of a single retinal ganglion cell therefore signals that the retinal image contains contrast at a particular location and spatial scale. The distinction between the off- and on-centre cells also allows the phase of the input signal to be encoded by the activity of the ganglion cells (De Valois & De Valois, 1990a).

2.2.2 The brain

The signals from the retinal ganglion cells are communicated to the lateral geniculate nucleus (LGN) first by the optic nerve and then by the optic tract. The properties of the receptive fields at this stage are largely similar to those of ganglion cells in the retina, albeit with some modification to their tuning (Bullier & Norton, 1979). After the LGN, signals travel via the optic radiation to the primary visual cortex (V1). Combination of inputs from cells in the LGN with spatially adjacent receptive fields allows the simple cells of V1 to have responses which are spatial frequency, phase and orientation tuned (Hubel & Wiesel, 1959, 1962). These cells are also the first to receive input from both eyes.

Within V1, the outputs from several simple cells converge onto complex cells (Hubel & Wiesel, 1962). Complex cells therefore inherit some of the tuning properties of their parent simple cells, however they also exhibit spatial and phase invariance (De Valois & De Valois, 1990c). That is, they respond to a stimulus of the appropriate scale and orientation when it is presented at any position within their (relatively large) receptive field. Following V1 the processing of the retinal image continues in the extrastriate areas, which (in part) feature cells with more complex receptive fields. Additional convergences and integrations over space and time occur to allow for the encoding of more complex information. There is a coarse separation of the processing of motion and form by distinct extra-striate regions of the cortex (Braddick, O'Brien, Wattam-Bell, Atkinson, & Turner, 2000), sometimes referred to as the "dorsal" and "ventral" streams respectively (Mishkin & Ungerleider, 1982; Goodale & Milner, 1992).

2.2.3 Inhomogeneities

The above gives a brief account of how signals from the retinal image are processed and combined in the early stages of the visual system, however these operations are not applied homogeneously across the visual field. The light coming into the eye is affected by optical factors that degrade the image at greater eccentricities. The retinal circuitry shows marked inhomogeneities in the densities of cones and ganglion cells, and in the degree of convergence in their outputs (Perry & Cowey, 1985; Curcio & Allen, 1990; Curcio et al., 1990; Anderson, Mullen,

& Hess, 1991). After the retina, there is further evidence of a preference for the preservation of information from the central visual field over that from more eccentric locations. The fovea represents only 0.005% of the input from the eye, but it takes up 10% of the cortical surface in V1 (Snowden, Thompson, & Troscianko, 2006). This rescaling of the projected image is referred to as the cortical magnification factor (Daniel & Whitteridge, 1961; De Valois & De Valois, 1990c). There is also a relative increase in the ratio of complex cells to simple cells in the periphery (Wilson & Sherman, 1976). These inhomogeneities in the neural architecture result in an impairment in performance for many visual tasks when they are performed away from the fovea (Strasburger, Rentschler, & Jüttner, 2011), and this introduces a possible confound into any experiment that presents stimuli which are extended over the visual field. In this thesis I develop a method that can be used to compensate for this confound for contrast detection tasks (Chapter 4).

2.3 Filters, channels and transducers in the visual system

2.3.1 Spatial filtering

The spatial frequency and orientation tuned simple cells in V1 can behave as a bank of filters*, deconstructing the input to the visual system into its Fourier components (Graham, 1981). It has been demonstrated that it is these components to which the cells respond (De Valois, De Valois, & Yund, 1979). The shape of the receptive field in the spatial domain resembles that of a Gabor patch (see Section 3.4.4), and these are frequently used as models of simple cells (Marčelja, 1980; Jones & Palmer, 1987; De Valois & De Valois, 1990c). Convolving a stimulus by a Gabor patch produces a filtered image, containing the spatial frequency and orientation components that fall within the passband defined by the dimensions of the Gabor. This passband can be visualised by taking the Fourier transform of the Gabor (Figure 3.3). In Fourier space it is represented by a two-dimensional Gaussian positioned at its centre frequency and orientation, with the radial and angular spread of that Gaussian corresponding to the spatial frequency and orientation bandwidths respectively.

A limitation of the Gabor patch model is that the response of simple cells to stimuli of different spatial frequencies is asymmetric on a linear spatial frequency axis and therefore does not correspond to the Gaussian profile of the Gabor model in the Fourier domain. The simple cell response *is* symmetric when plotted on a log spatial frequency axis however (De Valois, Albrecht, & Thorell, 1982), for this reason a log-Gabor patch is a more accurate model of the simple cell receptive field (see Section 3.4.5). In this thesis, the responses of simple cells with a particular spatial frequency and orientation tuning are simulated by convolving stimulus images with a log-Gabor. The value of each pixel in the output image indicates the response of

*Their receptive field structure would also allow simple cells to behave as edge detectors (Tolhurst, 1972; Shapley & Tolhurst, 1973; Marr & Hildreth, 1980).

a “detector” (a hypothetical neurone with its receptive field defined by the properties of the Gabor used in the filtering) at that location.

2.3.2 Spatial frequency and orientation channels

Groups of detectors (as described above) with the same spatial frequency and orientation tuning but positioned at different locations in the visual field can be described as belonging to a single “channel” (Graham, 1989c). Psychophysical methods are not precise enough to directly stimulate or measure the response of an individual detector, but a single channel *can* be investigated by presenting stimuli which contain a single dominant spatial frequency and orientation (this is the case for most of the studies presented in this thesis). The bandwidths of individual channels have been investigated psychophysically using masking (Campbell & Kulikowski, 1966; Phillips & Wilson, 1984), adaptation (Blakemore & Campbell, 1969) and subthreshold summation (Kulikowski, Abadi, & King-Smith, 1973) experiments.

2.3.3 Nonlinear transduction

Physiology studies indicate that the response of cortical neurones is nonlinear (Heeger, 1991, 1992; Tolhurst & Heeger, 1997; Carandini, Heeger, & Movshon, 1999). At low input magnitudes the responses of each detector are an accelerating function of the input, and at higher magnitudes these responses saturate. One suggested form of the accelerating nonlinearity is rectification followed by squaring (Heeger, 1991). Psychophysical evidence of nonlinear transduction is provided by threshold versus contrast functions from contrast discrimination experiments, which measure the smallest detectable increment to a pedestal stimuli of various magnitudes (Stromeyer & Klein, 1974; Legge & Foley, 1980; Foley & Legge, 1981; Kontsevich & Tyler, 1999). A linear transducer would predict that equal increments on a pedestal should be equally discriminable, however the data typically show a characteristic “dipper” function where small pedestal contrasts facilitate detection. This effect is predicted by power-law transduction of low contrasts.

A linear transducer also makes a characteristic prediction for the slope of the psychometric function. This is a measure of how an observers’ performance (expressed as proportion-correct) for detecting a stimulus should vary as a function of the stimulus magnitude. The psychometric function found in detection experiments is steeper than that predicted by a linear system, but consistent with a power-law transducer (Tanner & Swets, 1954; Swets, Tanner, & Birdsall, 1961; Lasley & Cohn, 1981). An accelerating nonlinear transducer is also supported by evidence from noise-masking studies (Lu & Dosher, 2008), and some findings from area summation experiments (see below).

2.3.4 Uncertainty

Consider a task where an observer is required to detect a single stimulus with prior knowledge of its spatial frequency, its orientation and the location in the visual field at which it will be presented. If performance is limited by internal noise in the response of each detector, the best strategy to adopt is to make the decision based only on detectors that would be expected to carry information about the stimulus. There are two conditions under which the observer cannot adopt this strategy. The first, known as "extrinsic uncertainty", is if there is not one single possible stimulus but several potential stimuli (Graham, Robson, & Nachmias, 1978; Davis & Graham, 1981; Davis, Kramer, & Graham, 1983; Graham, 1989a). In that case the observer must monitor the detectors that would respond to all possible stimuli on each trial, and so behave as if they are "uncertain" of the expected stimulus. The second condition under which the observer may behave as if they are uncertain is where they do have knowledge of the expected stimulus but are unable to apply that knowledge to modify their behaviour (Green, 1961; Tanner, 1961; Graham, 1989b). This is known as "intrinsic uncertainty".

Regardless of whether the uncertainty is caused by the experimental design (extrinsic) or limitations of the visual system (intrinsic), the same predictions result for how the slope of the psychometric function should be affected (Pelli, 1985). An observer operating under conditions of uncertainty would have steeper psychometric function slopes for detection, indistinguishable from those seen with a nonlinear transducer (Lasley & Cohn, 1981). Uncertainty can also predict the dipper shape of the threshold versus contrast function, if the pedestal has the effect of making the observer less uncertain of the properties of the expected stimulus (Tanner, 1961; Pelli, 1985). The similarity of the predictions made by models that involve intrinsic uncertainty and those that feature a nonlinear transducer makes a distinction between the two difficult to draw unless an experiment is designed to differentiate between them. This can be done by manipulating the level of extrinsic uncertainty (e.g. Meese & Summers, 2012).

2.4 Psychophysical theory

2.4.1 Sensory thresholds

One of the most basic investigations that can be made of the visual system is to measure thresholds (Green & Swets, 1966d). In general terms a "threshold" defines a point at which a system changes from one state to another. When used in psychophysics, "threshold" often refers to the magnitude of the minimum perceivable change in a stimulus along a particular dimension (e.g. contrast, orientation, size). It is thought that sensory systems are limited in this way (as opposed to being able to reliably detect even the smallest changes in a stimulus) due to the effects of random noise in the sensory system (Pelli, 1990). Here, this noise is assumed to be Gaussian and independent across space and time unless otherwise stated (Green

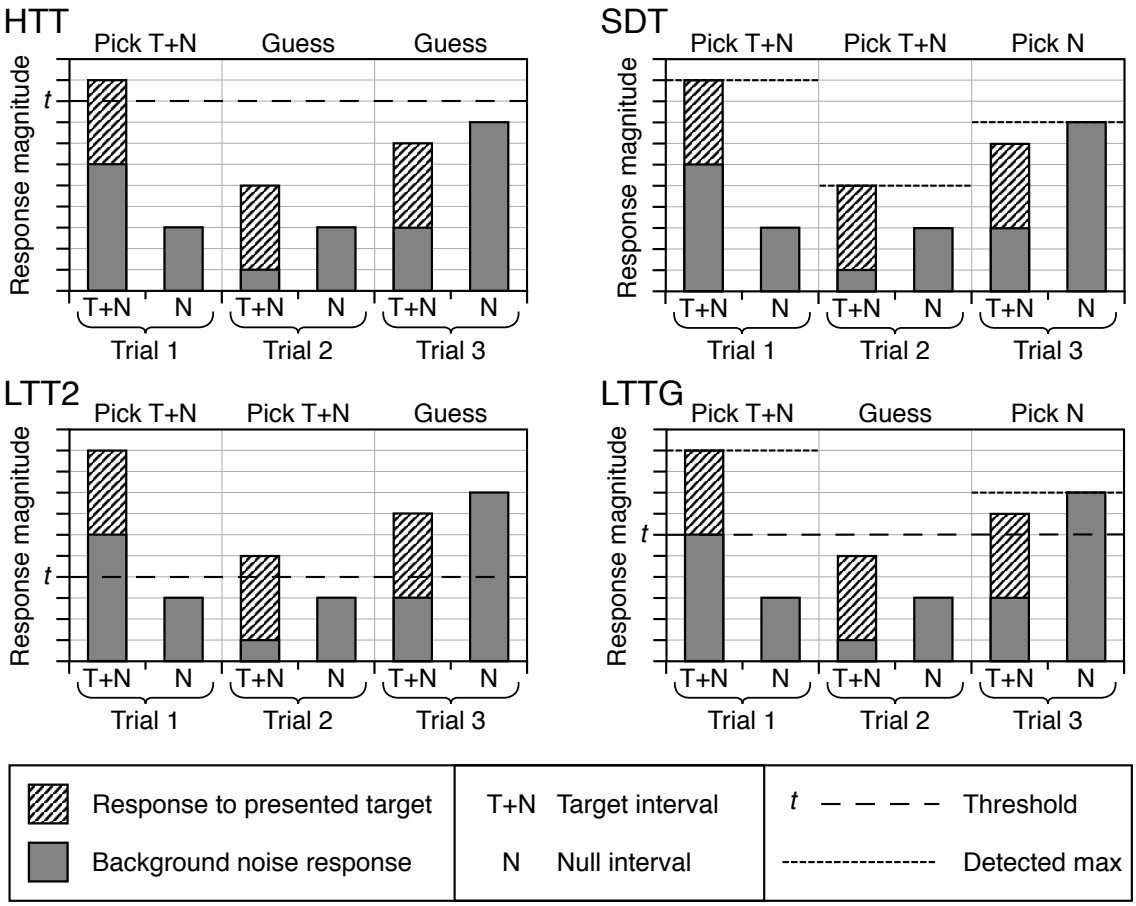

Figure 2.1: Graphs illustrating the behaviours of four detection theories: High Threshold Theory (HTT), Signal Detection Theory (SDT), Low Threshold Theory with a Two State response (LTT2), and Low Threshold Theory with a Graduated response (LTTG). They show how an observer behaving according to each model would decide to respond in three example trials of a psychophysical task. The task is to determine which of two intervals contained a signal that is targeted to activate a single detector, added to a background of internal noise in that detector. The noise levels on each trial and interval have been chosen to illustrate the different behaviours of the four models. When operating under the HTT or LTT2 model, the systems make their choice purely based on whether an interval (either the T+N or N) has exceeded an internal threshold, the difference between the two models being whether this threshold is set sufficiently high that it is never (or very rarely) activated by the internal noise alone. An observer operating according to the SDT model has access to a graduated response in each interval, so that the two values can be compared on each trial. The LTTG model also has access to this graduated response, but only for intervals where the internal (low) threshold has been surpassed.

& Swets, 1966a). As the noise causes the responses of observers to particular stimuli to be variable, it is typical to place the threshold at the magnitude where the change is perceptible with a particular probability (e.g. 50%).

It is not obvious however what this threshold level of performance refers to. The existence of a sensory threshold as a particular stimulus level with any special significance within psy-

chophysics has long been disputed (Swets, 1961; Corso, 1963; Green & Swets, 1966e). Frequently the term is used purely to refer to the stimulus level that results in a particular level of performance, and this shall be its meaning when used in the rest of this thesis unless otherwise indicated. Outlined below is a review of the four main models of "threshold" vision within the history of psychophysics (with the signal detection theory model removing the idea of a fixed threshold altogether). Figure 2.1 provides examples of the expected performance of sensory systems performing under the various proposed theories when engaged in a task where the observer must indicate which of two temporal intervals contained a signal against a background of internal noise (the noise levels having been chosen to illustrate the characteristic differences between these four models).

2.4.2 High threshold theory

Under high threshold theory (HTT; see Green & Swets, 1966e), it is proposed that detection is limited by fixed thresholds which are set sufficiently high (i.e. that require such a large magnitude of stimulation to be achieved) that the relevant detecting mechanism will be activated only when the stimulus to which it is tuned is presented, and will never (or very rarely) respond to mere background noise. The behaviour of the detecting mechanisms under HTT is binary, they either activate (indicating the presence of the stimulus) or do not (indicating its presumed absence). Information about the magnitude of the stimulation is not communicated through the system, though this may be inferred by the number of mechanisms that are activated or by the activation of different types of mechanism. The response that is then made to a stimulus (for example the answer to the question "Was the stimulus present or absent?") is determined by a higher level mechanism that monitors the outputs of these detectors (see Figure 2.1).

The application of HTT to the detection of threshold contrast has been challenged. In a task where the presence or absence of a signal has to be determined (a yes-no task) one of the consequences of HTT is that the the hit rate (the proportion of "stimulus present" responses when a stimulus is shown, in a task where the observer has to detect the presence of a stimulus) will be independent of the proportion of false alarms ("stimulus present" responses on occasions where no stimulus is shown), as false alarms should purely be due to guessing by the observer. It has been shown however that the false alarm and hit rates are not independent (Green & Swets, 1966e; Nachmias, 1981). In addition, in a task where observers had to rank four temporal intervals according to their perceived likelihood of containing a signal (which was only displayed in one interval), the second choices of those observers predicted the target interval with a greater probability than is possible under HTT (Swets et al., 1961).

2.4.3 Signal detection theory

Signal Detection Theory (SDT) was originally formulated to study the use of radar (Marcum, 1947), and was later developed and extended to analyse biological sensory systems (Peterson,

Birdsall, & Fox, 1954; Tanner & Swets, 1954; Green & Swets, 1966c). Under SDT there is no fixed threshold. The higher level decision mechanism has access to the continuous responses of each of the detecting mechanisms that it monitors. These responses are either compared against a criterion level for detection set by the observer (in a yes-no task), or against responses from another sampling interval (for tasks where the observer must determine which interval contains the target, as shown in Figure 2.1).

Unlike HTT, the implication of the criterion in SDT is that the false alarm rate should vary with the hit rate. This fits with human psychophysical data (Green & Swets, 1966e; Nachmias, 1981). It is also predicted under SDT that the observer should be able to make second choices in a four-alternative task at better than chance level. This prediction is accurate to data collected by Swets et al. (1961). Based on this evidence and these two candidate theories alone, HTT should be rejected in favour of SDT when considering the detection of threshold contrast.

2.4.4 Low threshold theories

There is also the possibility that performance is limited by a threshold which is regularly exceeded by the noise alone. Two of these Low Threshold Theories (LTT) are examined here. The first of these is similar to HTT, however the threshold is set low enough that it is exceeded by the noise on a significant number of trials. This Low Threshold Theory with a Two State Response (LTT2 in Figure 2.1; also known as multi-threshold two-state theory; see Green & Swets, 1966e) is incompatible with *a posteriori* probabilities of the presence of the stimulus calculated in rating-scale experiments (Nachmias & Steinman, 1963; Krantz, 1969).

The second theory is the Low Threshold Theory with a Graduated Response (LTTG in Figure 2.1, also known as single-threshold multi-state theory; see Swets et al., 1961; Green & Swets, 1966e). The evidence given above for SDT over HTT would also be compatible with LTTG (Green & Swets, 1966e). There is no reason however to favour LTTG over SDT for the detection of contrast, and the implementation of LTTG would involve adding an additional (difficult to measure) parameter to the SDT model. For reasons of parsimony, the assumption will be that there is no low threshold, and SDT will be used in this thesis.

2.5 The interpretation of area summation data

2.5.1 Combining signals from several detectors

Figure 2.1 presents examples of how an observer could use the responses from a single detector across two intervals to make a response in a psychophysical task, however in practice any stimulus presented to the visual system will activate many detectors (either within a single

channel or across multiple channels). The observer can then somehow combine the outputs of these detectors to improve their performance (e.g. detect weaker signals). In this thesis I investigate how signals are combined from detectors positioned across the visual field. The results from area summation experiments indicate that such processes occur, as larger stimuli tend to be detectable at a lower magnitude than smaller stimuli. Combinations of detection theory and hypotheses about neuronal architecture can be used to make predictions about the decrease in threshold expected for a particular increase in stimulus size.

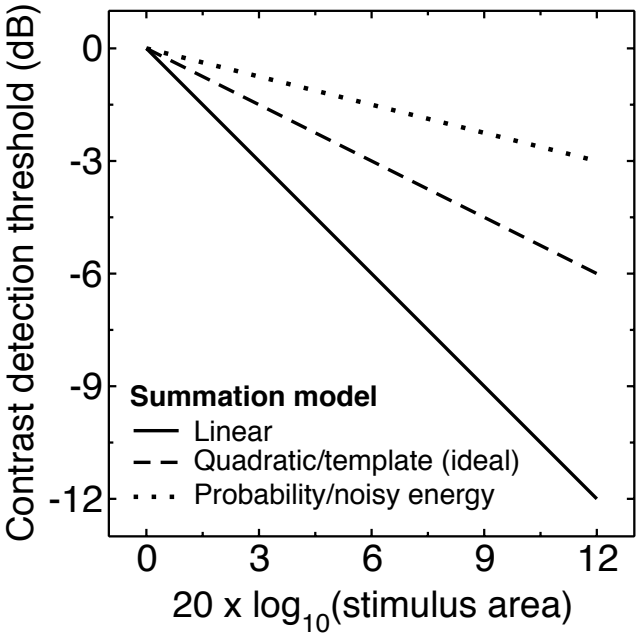

Figure 2.2: The summation slopes predicted by five different summation models, showing how detection threshold (in dB, see Section 3.2) declines as a function of stimulus area (expressed as $20 \times \log_{10}(\text{stimulus area})$ for ease of presentation). The linear summation prediction has a slope of -1. The quadratic and ideal summation predictions both have a slope of $-\frac{1}{2}$. The probability and noisy energy (ideal summation with a square-law transducer) predictions both have a slope of $-\frac{1}{4}$.

2.5.2 Linear summation

One theory of how signals are combined across the visual field is that a particular stimulus magnitude is required for detection, but that this same quantity can be spread out over space without affecting stimulus detectability. This proposal, called "Ricco's law" when applied to the summation of luminance increments (Graham, Brown, & Mote, 1939) but referred to as "linear summation" when applied to the summation of contrast here, predicts a linear slope with a gradient of -1 (see Figure 2.2) when plotting the logarithm of the threshold against the logarithm of the stimulus area (i.e. a doubling in the stimulus area should result in a halving of the threshold). This is the expected behaviour for summation within a single detecting mechanism if performance is limited by additive noise. This prediction also holds for other cases where the

standard deviation of the limiting noise remains constant for different stimulus sizes. For example, summation between two detecting mechanisms where their outputs are mandatorily added together to form the input to a single higher-level mechanism (summing the magnitude of the signal and the variances of the independent per-detector Gaussian noise).

2.5.3 Summation following nonlinear transduction

It is also possible to obtain shallower summation slopes when outputs are being combined across detectors. In the case where the outputs from each detector pass through an accelerating nonlinearity (for example, squaring; see Section 2.3.3) before being added together, less summation will be measured (these "quadratic summation" processes are of the sort required if the higher-level mechanisms they feed into are to represent stimulus energy; see Manahilov, Simpson, & McCulloch, 2001; Meese, 2010). A square-law transducer will result in a summation slope of $-\frac{1}{2}$ (see Figure 2.2), which has previously been reported under the name of "Piper's law" for luminance increments (Graham et al., 1939).

2.5.4 Ideal summation using templates

A summation slope of $-\frac{1}{2}$ can also be seen for linear detectors when responses are combined optimally across them. This "ideal summation" model assumes that each detector is affected by independent noise (which is the perfomance-limiting noise in the system), and that the observer knows which detectors will be activated by the stimulus on each trial (a blocked experimental design) and uses a matched template in order to only take the output from those detectors (Kersten, 1984; Tyler & Chen, 2000; Meese, 2010). By ignoring the noise from irrelevant detectors when detecting a smaller stimulus, thresholds for those conditions are improved leading to a shallowing of the summation slope. A $-\frac{1}{2}$ slope is also predicted in experiments that interleave several stimuli, provided that the observer has a template matched to each stimulus and is able to perform a max operation (weighted by the expected noise in each template) on each trial to pick the template that is most likely to be matched to the presented stimulus (Tyler & Chen, 2000).

The similarity of the summation slope predictions between the ideal and quadratic summation models means that they cannot be distinguished by measuring summation slopes alone without designing the experiment in such a way that divergent predictions are made. For example although the ideal summation model predicts the same summation slope for the blocked and interleaved designs described above, the absolute threshold prediction will be higher for the interleaved case (Meese & Summers, 2012) because of the trials on which the wrong template is selected by the max operation (based on the noise level in that template exceeding the sum of the signal and noise in the correct template). This is an effect of extrinsic uncertainty. A quadratic summation model would predict that performance in the blocked and interleaved conditions should be the same.

2.5.5 Probability summation under HTT

Area summation experiments typically find shallower summation slopes than those predicted by the linear, quadratic and ideal summation models described above. This has typically been explained as resulting from probabilistic summation between the outputs of independently-noisy detectors. These "probability summation" models were originally built on the framework provided by High Threshold Theory, describing the psychometric function as a Weibull function (Nachmias, 1981; Pelli, 1985). The probability of a stimulus being detected by any one of the detectors monitoring a location at which it appears is equal to 1 minus the probability that no detector was activated by the stimulus (Robson & Graham, 1981). Model predictions from this theory are usually approximated using Minkowski summation (see Section 5.4.2; Quick, 1974), where the slope (β) of a Weibull function fitted to the detection data is used to control the summation slope (that slope being $-\frac{1}{\beta}$, with a typical value of β being 4, see Figure 2.2).

2.5.6 Probability summation under SDT

The probability summation account was updated by Pelli (1985). The HTT assumption was removed, and instead the observer was said to be performing the summation tasks under conditions of intrinsic uncertainty (as HTT behaviour is approximated by a model observer operating under sufficiently high uncertainty). This model changes the theoretical interpretation of what processes that the visual system is engaged in, without changing the way in which these are represented mathematically (although there was a change from the mathematical models describing the exact performance of the observer to merely being close approximations to their performance). This account has been challenged for requiring an implausible degree of intrinsic uncertainty on the part of the observer (Tyler & Chen, 2000).

An extensive investigation of different models of probability summation under SDT was undertaken by Tyler and Chen (2000), and later extended by Meese and Summers (2012). The key result from the analyses performed was that probability summation under SDT is approximated by Minkowski summation with a fixed β of 4 (see Section 5.4.3), provided that the transducer is linear and that the observer is uncertain about which detectors carry information about the stimulus. As this value is typical of Weibull βs found in contrast detection experiments, this model also describes the data from area summation studies that are well-fitted by HTT probability summation models.

2.5.7 Combination models

Studies that measure contrast sensitivity to stimuli of various sizes find steeper summation slopes for small stimuli (Foley et al., 2007; Meese & Summers, 2007, 2012). Many contemporary models of area summation feature an initial filtering stage, allowing for linear summation

within the footprints of the filter elements (these can be analogous to simple cell receptive fields). This results in a summation slope of -1 for small stimuli, which then becomes shallower as stimuli are increased in size to be larger than the filter element.

It is also possible to achieve shallower summation slopes by combining different components in one model. A model that features both a square-law transducer (slope of $-\frac{1}{2}$) and ideal summation (slope of $-\frac{1}{2}$) would have a summation slope of $-\frac{1}{4}$ (and thus be difficult to distinguish from probability summation when thresholds alone are measured, see Figure 2.2; see also Meese, 2010; Meese & Summers, 2012). Similarly, a model that combined a square-law transducer (slope of $-\frac{1}{2}$) applied to each detector with probability summation (slope of $-\frac{1}{4}$) between those detectors would predict a summation slope of $-\frac{1}{8}$ (Meese & Summers, 2012).

2.6 Summation of contrast to threshold

2.6.1 Stimuli used in area summation experiments

Traditionally, stimuli used to investigate area summation in vision were luminance increments. These were either presented as bright discs or Gaussian blobs of various diameters (Graham et al., 1939; Barlow, 1958; Bijl, Koenderink, & Koenderink, 1993). However, following the recognition that stimuli are processed by the visual system in separate channels corresponding to different spatial scales (Section 2.3.2), experiments were then conducted using spatially bandpass stimuli such as windowed sine-wave gratings and Gabor patches (see Section 3.4). Gratings are superior to luminance increments due to the fact that varying the size of a luminance stimulus will shift its energy to lower spatial frequencies. This makes the area summation properties of data from those stimuli more difficult to interpret, as changes in the size of the stimulus are confounded with changes in its spectral properties. On the other hand, increasing the area of a contrast-defined stimulus such as a grating will cause it to activate a greater number of spatially distributed detectors tuned to the same spatial frequency. Further support for the use of gratings comes from experiments conducted using periodic stimuli that show a dependence of the summation effect on the number of cycles of the stimulus shown, rather than a reliance on absolute aperture size (Hoekstra, van der Groot, van den Brink, & Bilsen, 1974; Howell & Hess, 1978).

2.6.2 Previous results from luminance-defined stimuli

For luminance-increment stimuli, psychophysical data conform to Ricco's law when the stimulus's size is varied within an area which is relatively small. In the fovea, linear summation is found up to three minutes of arc (Graham et al., 1939), in the periphery this behaviour extends

over larger areas (0.4 - 0.5 deg^2 in Graham et al., 1939; Barlow, 1958). These data are consistent with linear summation within the smallest available receptive field of the observer at that eccentricity. Increasing the size of a luminance stimulus beyond this typically uncovers a range of sizes over which summation has a slope of $-\frac{1}{2}$ (following Piper's law; Barlow, 1958; Bijl et al., 1993). Ideal summation of this kind would be predicted if the observer had access to equally sensitive receptive fields at every spatial frequency and was certain about the stimulus being shown, as the entire stimulus could then be detected by a single receptive field matched to its extent. For larger areas summation typically continues at progressively lower rates until thresholds become entirely independent of stimulus size (Bijl et al., 1993).

2.6.3 Previous results from contrast-defined stimuli

Linear summation has been found for grating stimuli when their size is varied within the range expected to be detected by the receptive field of a single simple cell (Legge, 1978; Kersten, 1984; Polat & Tyler, 1999; Foley et al., 2007). This level of summation is also seen between a "full" grating stimulus and a contrast modulated version ("Swiss cheese" or "Battenberg") which has been multiplied by a checkerboard plaid to introduce small "checks" of full and zero contrast (see Section 3.4.3 and 3.4.6; Meese & Summers, 2007; Meese, 2010; Baker & Meese, 2011). These linear short-range summation effects are predicted by any model which includes a linear filtering stage (for example convolution by Gabor approximations to simple cell receptive fields). As the stimulus is increased in size the summation slope transitions from -1 (linear summation) to a shallower gradient. This usually involves a region with a slope of $-\frac{1}{2}$, though this may be an artefact of a transition between a slope of -1 and a slope of $-\frac{1}{4}$, rather than representing the operation of summation processes that produce a slope with this gradient.

In the fovea, investigations of summation confound effects of stimulus size with those of the inhomogeneity in contrast sensitivity across the visual field. This will lead to studies tending to underestimate the level of summation between the outputs of the detectors. Nevertheless, Kersten (1984) found summation slopes of $-\frac{1}{2}$ over four grating cycles in the fovea. Other studies have explained foveal data by incorporating a model of the sensitivity inhomogeneity into their summation models, these studies typically find summation consistent with that which would produce slopes of $-\frac{1}{4}$, if the inhomogeneity were not present (Robson & Graham, 1981; Watson & Ahumada, 2005; Meese & Summers, 2007, 2012). Originally these data were presented as evidence for probability summation (Robson & Graham, 1981), however the contemporary "noisy energy" model (which combines square law transduction with a template matched to the stimulus extent) also accounts for these data (Meese & Summers, 2012).

The effects of the inhomogeneity in contrast sensitivity across the visual field can be avoided by presenting stimuli in the periphery, as there are areas where its effect is very small (Robson & Graham, 1981). Results from summation experiments conducted in the periphery are inconsistent however. Mayer and Tyler (1986) found linear summation (slope of -1) over more

than eight stimulus cycles in the periphery (much larger than the estimated size of simple cell receptive fields), however the results of this study have not been replicated elsewhere in the literature. Some investigations have found spatially extensive summation with a slope of $-\frac{1}{2}$ (Manahilov et al., 2001; Meese & Hess, 2007), whereas other studies have found only fourth-root summation when testing over a wider area (Robson & Graham, 1981). It is not clear why these differences in summation slope have been found. The steeper slopes for smaller stimuli in the periphery could be explained in the combination model if one of the two components shallowing the slope became inactive under those conditions, either by linearisation of the transducer or by a breakdown in the template stage (resulting in mandatory pooling over an inefficiently large area).

2.6.4 Conclusions from previous contrast detection threshold studies

From previous data, the two models that provide good accounts of the data are the probability summation model and the noisy energy model that combines square-law transduction with a template stage (with both models including initial linear filtering stages). Additional evidence from the empirical slopes of the psychometric functions compared against those predicted by the models favours the noisy energy combination model (Meese & Summers, 2009, 2012). Models of extended summation in the fovea have however always relied on their accounts of the inhomogeneity in contrast sensitivity as a major determinant of the predictions that they make. In addition, the inconsistent results from summation experiments conducted in the periphery make assessments of behaviour here difficult to conduct conclusively without further studies (as data could be selected from the literature to fit any viable model). These outstanding issues are addressed in Chapter 6 and Chapter 7 of this thesis.

2.6.5 Area summation in contrast discrimination studies

Early studies investigating the summation of *suprathreshold* contrast using the contrast discrimination method found a lack of area summation above threshold when a target of increasing size was placed on a pedestal matched to its extent (Legge & Foley, 1980). These findings are compatible with the probability summation hypothesis, provided noise becomes correlated between detectors above threshold. By manipulating target and pedestal extent separately however, studies have found greater levels of summation (Bonneh & Sagi, 1999; Meese, Hess, & Williams, 2005). By measuring contrast discrimination thresholds for Swiss cheese gratings (see Section 3.4.3) added to a full grating pedestal, Meese and Summers (2007) found substantial summation inconsistent with probability summation, but consistent with the physiological combination of transduced signals (combined with a suppressive contrast gain control; see Heeger, 1992; Foley, 1994).

2.7 Signal combination over area in other domains

2.7.1 Summation of motion

Analogous to the area summation of local contrast, motion information can be combined over space between individual moving elements to form a percept of global motion. Typically these studies are carried out by having an observer identify the direction of coherent motion hidden within an array of elements that are independently moving in random directions. As in the contrast summation case, it is expected that performance should improve as the number of samples increases. Studies have found the summation of motion to be spatially extensive, over 3.5-9 degrees (Downing & Movshon, 1989; Watamaniuk & Sekuler, 1992; Ledgeway, McGraw, & Simmers, 2011). The process by which the information is summed across individual detectors is proposed to be either a winner-take-all (max) operation over channels tuned to different motion directions, or a maximum likelihood estimation[†] over those channels (Webb, Ledgeway, & McGraw, 2007). However, the improvement in human performance with increasing stimulus area underperforms these model predictions (Watamaniuk, 1993).

2.7.2 Summation of orientation

The orientations of individual elements in a pattern can be combined over space in order to make judgements about the resulting texture (Dakin & Watt, 1997). Studies of how this is performed are often similar to those used for motion tasks, where a certain proportion of a set of otherwise randomly oriented elements are assigned the target orientation. It is proposed that the information is combined by either a vector average (as occurs for crowded stimuli; see Parkes, Lund, Angelucci, Solomon, & Morgan, 2001), a winner-takes-all (max) operation (Husk, Huang, & Hess, 2012), or a maximum likelihood estimation, depending on the stimulus conditions and the task set to the observer (Webb, Ledgeway, & McGraw, 2010).

Investigations of the summation of orientation information over area have found conflicting results. A "global form" detection experiment using glass pattern stimuli has found summation for concentric patterns but not for patterns of parallel elements (Wilson, Wilkinson, & Asaad, 1997). An absence of summation for parallel stimuli has also been found in an orientation discrimination experiment using arrays of narrowband elements (Husk et al., 2012). Equivalent noise analysis has shown however that the integration of orientation signals is governed by an information limit, and that performance improves as the size of the stimulus increased up to a maximum of 7 degrees (for 4 c/deg Gabors; see Dakin, 2001). A study carried out using

[†] The maximum likelihood estimation (MLE) strategy uses the response of the system to the stimulus to compute the likelihood that each possible stimulus was presented. The decision can then be made based on which stimulus was most likely to have been presented. This is the method that the ideal observer would use for this task, as it bases the decision on the likelihood ratio (Green & Swets, 1966b).

filtered noise patterns found less-than ideal summation for coherent local signals up to a maximum size of 2 degrees (Jones, Anderson, & Murphy, 2003).

Some of the factors that complicate the study of the integration of orientation signals over area mirror those encountered in studies of the contrast area summation. For example, management of extent of the stimulated region and the confounding effect of inhomogeneity in sensitivity (Jones et al., 2003) affect both types of study. An investigation of the summation of orientation signals over space, inspired contrast detection studies, is carried out in Chapter 9.

2.8 Amblyopia

2.8.1 Aetiology

One useful strategy to investigate the processing that occurs in the visual system is to look at situations where it functions abnormally. One such condition which is investigated as part of this thesis is amblyopia. This is a neural condition where an impoverished input from one eye during development (usually due to either strabismus or anisometropia) results in an long-term impairment in vision when that eye is used, even after the optical defect is corrected. Evidence suggests that the locus of this effect is largely cortical (Blakemore & Vital-Durand, 1986; Anderson, Holliday, & Harding, 1999; Barnes, Hess, Dumoulin, Achtman, & Pike, 2001), and that the size of the neural impairment depends on how poor the signal is from the impoverished eye during development (Smith, Hung, & Harwerth, 2000). It is thought that the binocular cells fail to develop properly in the cortex when the input from one of the eyes is weak or absent (Hubel & Wiesel, 1970), and that the input from the deprived eye must be restored before the end of some critical period if normal function is to be obtained (Blakemore & van Sluyters, 1974).

2.8.2 Effects on performance in visual tasks

Amblyopic observers using their affected eyes exhibit poor behaviour in many tasks, not limited to those requiring binocularity or high resolution acuity (McKee, Levi, & Movshon, 2003). The contrast sensitivity functions of amblyopes can be profoundly affected across the spatial frequency range (Hess & Howell, 1977; Levi & Harwerth, 1977). Noise-masking studies have found performance in detection tasks to be limited by a greater internal noise in amblyopes (Levi & Klein, 2008). Above threshold, veridical perception of grating stimuli is marred by spatial frequency specific distortions and scotomas (Hess et al., 1978). Performance on global motion tasks is impaired (Simmers, Ledgeway, Hess, & McGraw, 2003), as it is for orientation discrimination (Skottun, Bradley, & Freeman, 1986). The deficit that amblyopes show for tasks involving the combination of signals over space is investigated in this thesis by looking at spatial summation in amblyopes (see Chapter 8).

2.9 Conclusion

The combination of signals over space is a necessary stage in visual processing. One expected consequence of this signal combination is an improvement in performance for the detection of stimuli as they increase in size. Although the existence of a performance advantage for larger stimuli has been recognised in several domains, quantitative accounts of these improvements are still developing. In this thesis, novel stimuli and experimental methods are used to uncover new information about the processes underlying the integration of contrast signals over an increasing stimulus area. These methods are also applied to study the integration of coherent orientation signals across the visual field.

CHAPTER 3

General methods

3.1 Introduction

This chapter outlines the measurement conventions, equipment, stimulus paradigms, experimental techniques and methods of analysis which are used in this thesis.

3.2 Measures

3.2.1 Visual angle

Distances subtended by objects in the visual field are expressed in degrees of visual angle. To distinguish this from the orientation of a stimulus or filter (denoted by $^\circ$, see below) this is always presented as the full word "degree" or its abbreviation "deg". 1 degree of visual angle corresponds to approximately 0.3 mm on the retina (De Valois & De Valois, 1990b).

3.2.2 Spatial frequency

The spatial frequency of a sinusoidal or other periodic stimulus (see Section 3.4) is expressed as the number of cycles per degree of visual angle (c/deg).

3.2.3 Orientation

The orientation of a stimulus or a filter is referred to throughout this work by the use of a degree symbol ($^\circ$). An upright vertical stimulus will have an orientation of 0°, and this increases as it is rotated clockwise in the spatial domain.

3.2.4 Contrast

An image's contrast is the magnitude of its luminance modulation. There are several different metrics that are used to quantify contrast. The most commonly used is Michelson contrast which is defined as

$$c_{\text{Michelson}} = \frac{L_{\text{max}} - L_{\text{min}}}{L_{\text{max}} + L_{\text{min}}}, \tag{3.1}$$

where L_{max} and L_{min} are the maximum and minimum luminances in the image.

For stimuli whose mean luminance is not the mid-point of their minimum and maximum luminances, delta-contrast is typically used in order to make the calculated contrast independent of the phase of the stimulus (Section 3.4.1). Delta-contrast is defined as

$$c_{\text{delta}} = \frac{\max\left(|\mathbf{L} - L_{\text{mean}}|\right)}{L_{\text{mean}}}, \tag{3.2}$$

where L_{mean} is the mean luminance of the image ($\mathbf{L}$) and $\max\left(|\mathbf{L} - L_{\text{mean}}|\right)$ is the largest deviation from that value in the image. For any single stimulus, delta-contrast is linearly related to Michelson contrast. When not stated otherwise, "contrast" is intended as delta-contrast throughout this thesis.

Frequently it is more convenient to express contrast logarithmically. Contrast in dB re 1% is defined as

$$c_{\text{dB}} = 20\log_{10}(100 \cdot c_{\text{delta}}). \tag{3.3}$$

When not stated otherwise, "log-contrast" is intended as contrast expressed in dB re 1% delta-contrast throughout this thesis.

3.2.5 Size

For continuous grating stimuli, whether circular or rectangular, the dimensions W are derived from the full width at half magnitude (FWHM*) throughout this work. For stimuli with hard edges this is simply the width or height of the stimulus, however for those with blurred edges (for example windowed by a raised cosine function, see Section 3.4.1) the effective size of the stimulus extends beyond the plateau. The sizes of Gabor stimuli (see Section 3.4.4) are also expressed in this way when they are to be compared against those of gratings.

The radius of a circular stimulus, or the width and height of a rectangular stimulus, are then used to calculate its area. Here this is usually presented as $20 \times \log_{10}(\text{area})$, as plotting stimulus area in this way against detection thresholds expressed in dB gives an easily interpretable meaning to linear slopes of different gradients (for example, summation within a linear system would give a slope of -1, see Section 2.6.1).

*This measure is sometimes also referred to as the "full width at half height" (FWHH).

3.2.6 Signal-to-noise ratio

The signal-to-noise ratio (SNR) in a system is defined as the amplitude of the signal (μ) divided by the standard deviation (σ) of the noise affecting that signal

$$\text{SNR} = \frac{\mu}{\sigma}. \tag{3.4}$$

To find the SNR at the decision mechanism for a model involving the linear sum over several independent detectors (e.g. Section 5.3.2), the signal level in each channel is added. The combined signal is then divided by the standard deviation of the combined noise from these detectors (calculated by summing the variances).

3.2.7 Root-mean-square error

In this thesis, the differences between human behavioural data and model predictions that describe expected behaviour on that task are reduced to a single number by calculating the root-mean-square error (RMS error, or simply RMSe) between the data and the model

$$\text{RMSe} = \sqrt{\frac{\sum_{i=1}^{n} (\text{data}_i - \text{model}_i)^2}{n}}, \tag{3.5}$$

where data_i is the ith data point, model_i is the model prediction for that data point, and n is the number of data points.

3.3 Equipment

3.3.1 CRT monitors

In all experiments, stimuli were presented on cathode ray tube (CRT) monitors. These were gamma-corrected according to measurements taken with photometers in order that a linear series of requested increments to the luminance of an individual pixel in software would result in a linear series of increases in the luminance of that pixel on the monitor (Poynton, 1998; García-Pérez & Peli, 2001). All experiments were performed in a darkened room, such that the monitor would be the dominant source of light to the observer's visual system.

3.3.2 Stimulus presentation

Most experiments (Chapters 4, 6 to 8) used a Cambridge Research Systems (CRS) ViSaGe visual stimulus generator to store and display the stimuli. The orientation discrimination ex-

periment carried out in Chapter 9 used only the built-in NVIDIA GeForce 9600M of a Macbook Pro, as the precise presentation of threshold-level contrasts was not required here (as opposed to in all other experiments, where this was required).

3.3.3 Software

Most of the experiments (Chapters 4, 6 to 8) were conducted using Liberator. This is a bespoke application for running psychophysical experiments developed by Robert Summers and Tim Meese at Aston University. Delphi was used to write code for additional functionality where required. The experiment presented in Chapter 9 was conducted using Psychtoolbox (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007) running under MATLAB version 2010a (The Mathworks Inc.; Natick, Massachusetts).

3.4 Stimuli

3.4.1 Gratings

Sine-wave gratings are widely-used stimuli for studying visual processing. They are generated by sinusoidally modulating the luminance of a light source (typically a computer monitor) in one dimension, for example

$$\mathbf{L}[x, y] = c \cdot \sin(x - x_0), \tag{3.6}$$

where c is the amplitude (equivalent to delta or Michelson contrast for gratings) and x_0 is the starting position (the location of the initial zero-crossing for a sine-wave). See panels a) and d) of Figure 3.1 for an example of a grating stimulus.

The popularity of gratings as stimuli stems from two sources. Firstly, according to the Fourier theorem any image can be broken down into the sum of a series of 2D sine-wave gratings of different amplitudes, frequencies, orientations and phases. This makes them a good basic component for the study of vision. Secondly, physiological and psychophysical evidence both indicate that the visual system performs filtering operations similar to a local Fourier analysis at each location in the image (De Valois & De Valois, 1990b).

Gratings of arbitrary contrast (c), orientation (θ), frequency (f) and phase (ϕ) can be generated by setting

$$u = 2\pi f \cos\theta, \tag{3.7}$$

$$v = 2\pi f \sin\theta, \tag{3.8}$$

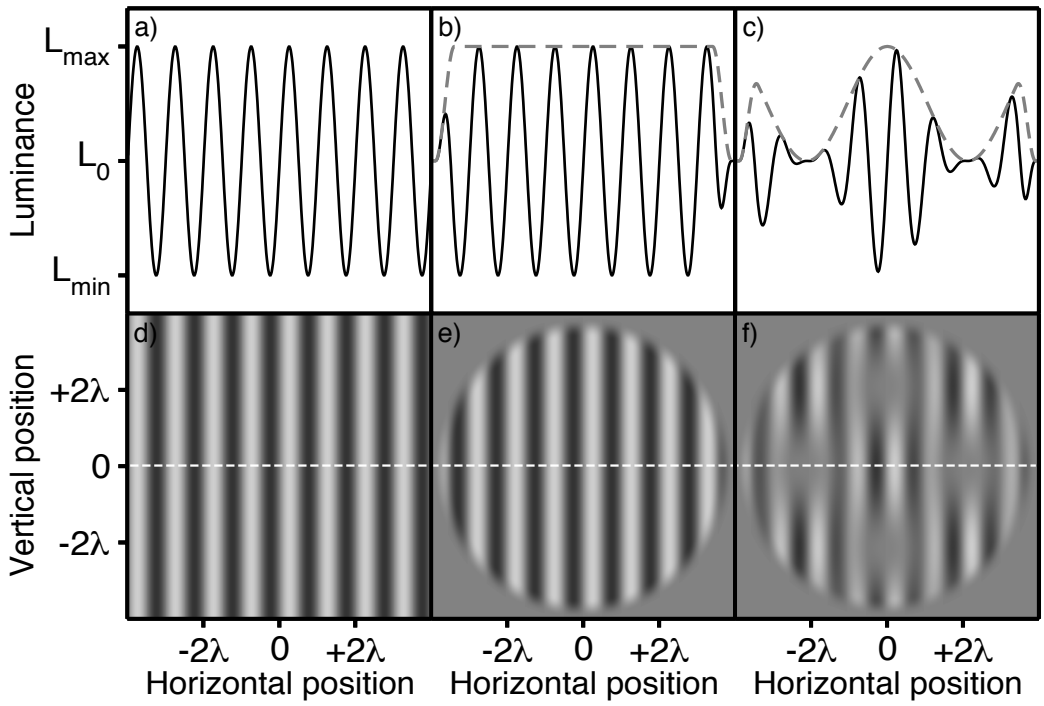

Figure 3.1: Horizontal cross-sections (a-c) through the centres of example stimuli (d-f). Black lines in panels a) to c) show the stimulus luminance (varying from its minimum at L_{min} to its maximum at L_{max} about the mean luminance L_0), the dashed grey lines show the envelope that is being applied to the carrier grating. Dashed white lines in panels d) to f) show where the cross section is taken. Panels a) and d) show a vertical sine-wave grating stimulus with wavelength λ. Panels b) and e) show this sine-wave grating windowed by a raised-cosine envelope with a plateau width of 6λ and a cosine section width of $\frac{1}{2}\lambda$. Panels c) and f) show that same windowed grating with "Swiss cheese" modulation. The modulator wavelength is 3.2λ, giving 1.6 cycles of grating per modulator "check".

and

$$\mathbf{L}[x, y] = c \cdot \sin\left[u(x - x_0) + v(y - y_0) + \phi\right]. \tag{3.9}$$

Grating patches of different sizes and shapes can be generated by multiplying the sinusoidal signal by an envelope. Typical shapes used include square, rectangular, and circular gratings. The edges of these envelopes are not usually sharp however as this generates energy at many additional spatial frequencies, the edges of the envelope can be blurred to ameliorate this.

3.4.2 Raised-cosine envelopes

For the studies presented here, gratings were presented within envelopes that declined from a gain of unity to a gain of zero in the manner of a single half-cycle of a cosine function (referred to as a "raised-cosine" envelope). The wavelength of this cosine function (λ) is twice the width over which the function needs to decline.

For circular stimuli this is calculated as

$$\lambda = 2(r_2 - r_1), \tag{3.10}$$

where r_1 is half the width of the stimulus plateau and r_2 is the half the width at which the envelope reaches zero. The envelope is then defined as

$$\mathbf{E}[x, y] = \begin{cases} 1 & \text{for } \sqrt{x^2 + y^2} < r_1 \\ \frac{1}{2}\langle 1 + \cos\left[\frac{\pi}{\lambda}(\sqrt{x^2 + y^2} - r_1)\right]\rangle & \text{for } r_1 < \sqrt{x^2 + y^2} < r_2 \\ 0 & \text{for } \sqrt{x^2 + y^2} > r_2 \end{cases}. \tag{3.11}$$

This makes the full width at half magnitude

$$W = r_1 + r_2, \tag{3.12}$$

and the nominal stimulus area

$$A = \pi \left(\frac{W}{2}\right)^2. \tag{3.13}$$

See panels b) and e) of Figure 3.1 for an example of a grating stimulus windowed by a circular raised-cosine envelope.

A rectangular stimulus may have a width that is different from its height. The start and end points must be defined separately in the two dimensions (w_{x1} and w_{x2} for the horizontal, w_{y1} and w_{y2} for the vertical), and the widths of the raised cosine sections may be different. The wavelengths for the cosine sections are calculated separately for the horizontal component (the shape of the decline in contrast along the x axis between the locations w_{x1} and w_{x2} being that of a cosine function with wavelength λ_x)

$$\lambda_x = 2(w_{x2} - w_{x1}), \tag{3.14}$$

and vertical component (the shape of the decline in contrast along the y axis between the locations w_{y1} and w_{y2} from y_0 being that of a cosine function with wavelength λ_y)

$$\lambda_y = 2(w_{y2} - w_{y1}). \tag{3.15}$$

Envelopes are defined separately for the horizontal and vertical directions

$$\mathbf{E}_x[x, y] = \begin{cases} 1 & \text{for } x < w_{x1} \\ \frac{1}{2}\langle 1 + \cos\left[\frac{\pi}{\lambda_x}(x - w_{x1})\right]\rangle & \text{for } w_{x1} < x < w_{x2} \\ 0 & \text{for } x > w_{x2} \end{cases}, \tag{3.16}$$

$$\mathbf{E}_y[x,y] = \begin{cases} 1 & \text{for } y < w_{y1} \\ \frac{1}{2}\langle 1 + \cos\left[\frac{\pi}{\lambda_y}(y - w_{y1})\right]\rangle & \text{for } w_{y1} < y < w_{y2} \\ 0 & \text{for } x > w_{y2} \end{cases}, \tag{3.17}$$

and then multiplied to form the rectangular envelope

$$\mathbf{E}[x,y] = \mathbf{E}_x[x,y] \cdot \mathbf{E}_y[x,y]. \tag{3.18}$$

The full width at half magnitude for the horizontal axis is given by

$$W_x = w_{x1} + w_{x2}, \tag{3.19}$$

and for the vertical axis

$$W_y = w_{y1} + w_{y2}, \tag{3.20}$$

and from these the nominal area† can be calculated

$$A = W_x W_y. \tag{3.21}$$

The envelope defined by Equation 3.11 or Equation 3.18 is applied to the grating (defined in Equation 3.9) as follows

$$\mathbf{L}[x,y] = c \cdot \mathbf{E}[x,y] \cdot \sin\left[u(x - x_0) + v(y - y_0) + \phi\right]. \tag{3.22}$$

3.4.3 Swiss cheeses

Modulating the contrast of a sine-grating by a raised plaid (scaled from 0 to 1) formed from two orthogonal gratings creates stimuli with checkerboard patterns of unity and zero contrast, with smooth modulation between the two. These stimuli allow the area of the stimulus containing contrast (signal area) to be varied whilst keeping the stimulus extent constant.

This "Swiss cheese" modulation is of the form

$$\begin{aligned} \mathbf{M}[x,y] = \frac{1}{2} + \frac{1}{4}\langle &\cos\left[2\pi f_m(x\cos\theta_m + y\sin\theta_m) + \phi_m\right] + \ldots \\ &\cos\left[2\pi f_m(x\cos -\theta_m + y\sin -\theta_m) + \phi_m\right]\rangle, \end{aligned} \tag{3.23}$$

where f_m, θ_m and ϕ_m are the frequency, orientation and phase of the modulation. This can be applied to a raised-cosine enveloped grating to produce the Swiss cheese grating stimulus

$$\mathbf{L}[x,y] = c \cdot \mathbf{E}[x,y] \cdot \mathbf{M}[x,y] \cdot \sin\left[u(x - x_0) + v(y - y_0) + \phi\right]. \tag{3.24}$$

†This will slightly overestimate the area of the envelope which is above half-magnitude due to the combined attenuation at the corners from the x and y components. This inaccuracy is negligible for small values of λ however.

See panels c) and f) of Figure 3.1 for an example of a Swiss cheese modulated grating stimulus windowed by a circular raised-cosine envelope.

The effective signal area of a Swiss cheese stimulus depends on the relationship between the diameter and the modulation frequency. Where values for the signal area are presented they are calculated as the area of the stimulus where the envelope is above half its maximum value (equivalent to calculating its area from the dimensions at half-magnitude).

3.4.4 Gabor patches

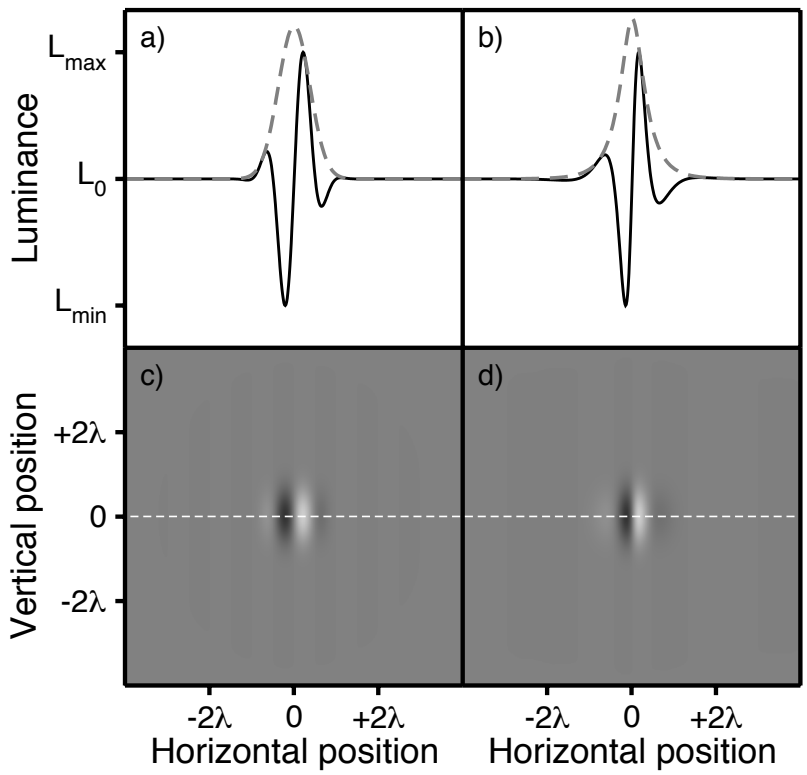

Figure 3.2: Panel c) shows a vertical Gabor stimulus with wavelength λ, an orientation bandwidth of $\pm 25°$ and a spatial frequency bandwidth of 1.6 octaves. Panel a) shows a cross-section through its horizontal centre (the dashed white line in panel c). The luminance profile of the stimulus (varying from its minimum at L_{min} to its maximum at L_{max} about the mean luminance L_0) is plotted in black, the envelope in dashed grey. Panels b) and d) show a log-Gabor stimulus with the same wavelength and bandwidths, and its cross-section.

Another popular stimulus for research into visual perception is the Gabor patch. This is a sine-wave grating windowed by a two dimensional Gaussian envelope (see Figure 3.2c; Gabor, 1946; Graham, 1989d). A Gaussian envelope with its major and minor axes along the cardinal orientations is defined as

$$\mathbf{E}[x, y] = \exp\left(\frac{-x^2}{2\sigma_x^2}\right) \cdot \exp\left(\frac{-y^2}{2\sigma_y^2}\right), \tag{3.25}$$

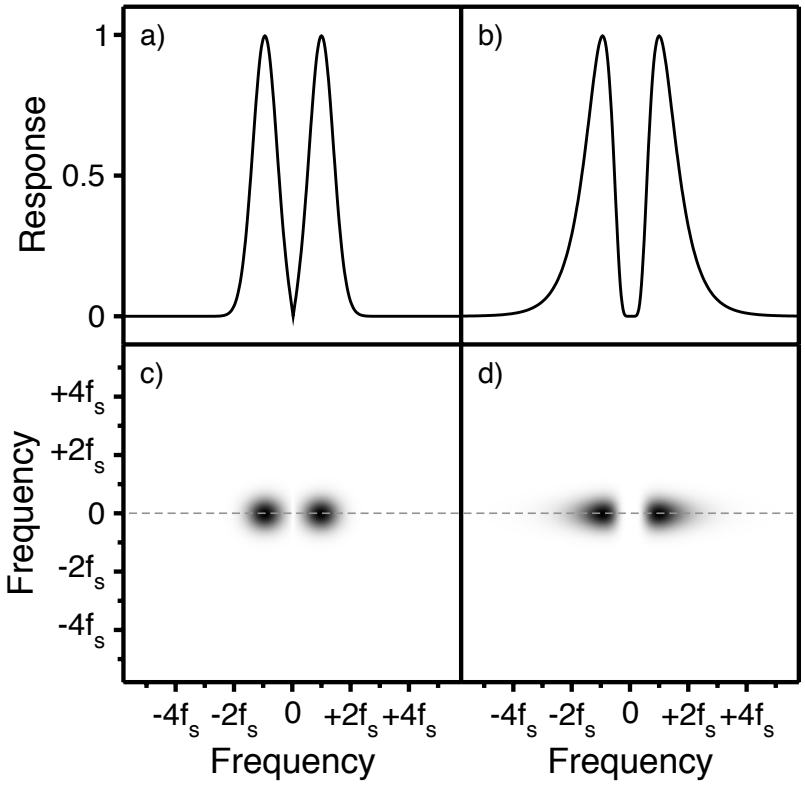

Figure 3.3: Panels c) and d) show the Fourier transforms of the Gabor and log-Gabor stimuli in panels c) and d) of Figure 3.2. The spatial frequency passbands of these stimuli at their preferred orientations (along the grey dashed line) are shown in panels a) and b). These are plotted on linear axes expressed in terms of f_s, which is the preferred spatial frequency of the Gabor and log-Gabor patches.

where σ_x and σ_y are the standard deviations of the Gaussian function along the horizontal and vertical axes. This envelope is applied to the grating (defined in Equation 3.9) as follows

$$\mathbf{L}[x, y] = c \cdot \mathbf{E}[x, y] \cdot \sin\left[u(x - x_0) + v(y - y_0) + \phi\right]. \tag{3.26}$$

For Gabor stimuli, the full width at half magnitude along an axis can be calculated from the standard deviation (σ) of the Gaussian envelope for that axis (Graham, 1989d) as

$$W = 2.35\sigma. \tag{3.27}$$

The Fourier transform of a Gabor patch is also the product of two Gaussians, positioned at the Gabor's preferred spatial frequency (see Figure 3.3c). The widths of these Gaussians are inversely proportional to the widths of the Gaussian envelope in the spatial domain. Gabor patches can also be used to filter images by convolution in the spatial domain, or multiplication in the Fourier domain. It has been demonstrated that local filtering operations of this kind are performed by the visual system (see Section 2.3.1; De Valois et al., 1979).

3.4.5 Log-Gabor patches

Log-Gabor patches share some of their properties with Gabor patches. However, a log-Gabor is defined in Fourier-space as the product of a one-dimensional log-Gaussian of spatial frequency and an orthogonal one-dimensional Gaussian of orientation. MATLAB code to produce log-Gabor patches is provided in Section B.1. An example of a log-Gabor patch is shown in panels b) and d) of Figure 3.2; its Fourier transform is shown in panels b) and d) of Figure 3.3.

Log-Gabor patches are used as models of the receptive fields of simple cells in V1 (see Section 2.2.2). Those used in this thesis have bandwidths of 1.6 octaves and $\pm 25^\circ$ respectively. These values are typical of those used previously in the literature, and compatible with those found in neurophysiological investigations (De Valois et al., 1982; De Valois & De Valois, 1990c; Meese & Summers, 2007; Meese, 2010).

Log-Gabors are used as stimuli in experiments where the intention is to provide simple cells with their ideal input. An additional practical advantage of using log-Gabors as stimuli is that they are DC-balanced (the sum of the signed pixelwise deviations from mean luminance is equal to zero) in any phase, and so have no effect on the mean luminance of the area of the display in which they are shown.

3.4.6 Battenberg patterns

The "Battenberg" stimulus was originally used by Meese (2010) to investigate the summation of contrast over area. It is a pattern made up of a square array of repeated elements, each one being a single cycle of sine wave grating at the target spatial frequency multiplied by an orthogonal half-cycle of cosine phase grating at half that spatial frequency. These micropatterns can be arranged in different configurations to flexibly generate smooth-edged grating-like stimuli whose signal area scales linearly with the number of elements used. Removing a subset of the elements (by setting their contrast to zero) allows for the creation of patterns which have a lower contrast over area, but the same extent as a "full" stimulus.

Entire Battenberg arrays can be generated at once by modulating the contrast of a full grating. The orthogonal rectified cosine grating (to produce an array of square Battenberg elements) is defined relative to the carrier grating (Equation 3.9) by the following equations

$$u_b = \pi f \cos(\theta - 90), \tag{3.28}$$

$$v_b = \pi f \sin(\theta - 90), \tag{3.29}$$

and

$$\mathbf{E}[x, y] = |\sin[u_b(x - x_0) + v_b(y - y_0) + \phi_b]|\,. \tag{3.30}$$

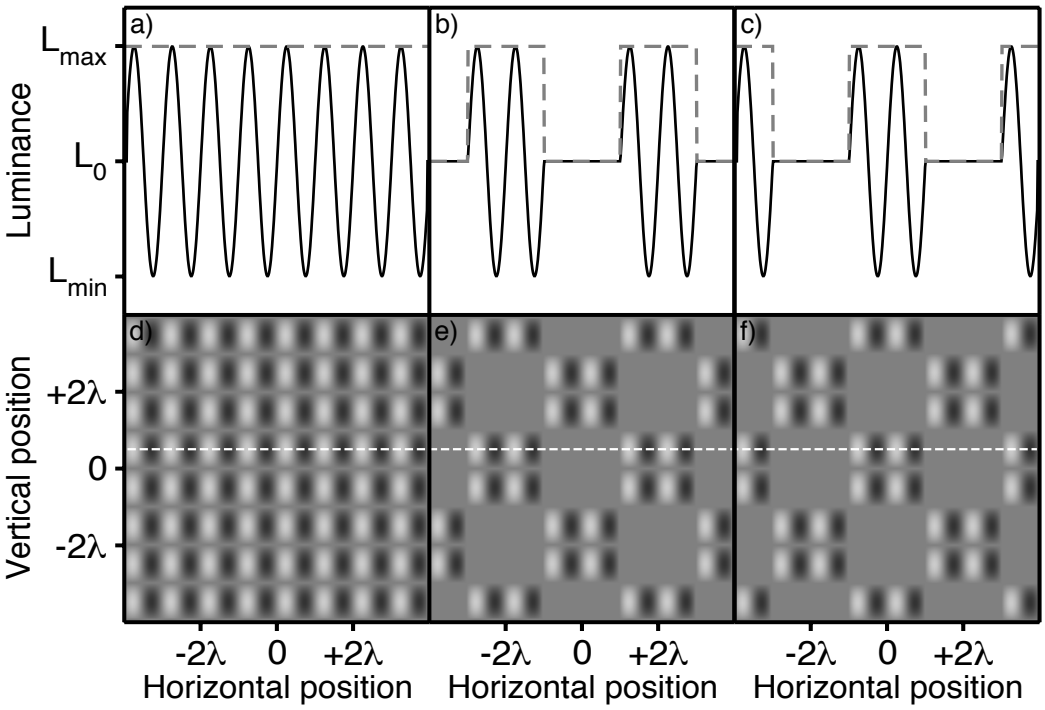

Figure 3.4: Horizontal cross-sections (a-c) through example "Battenberg" stimuli (d-f). Solid black lines in panels a) to c) show the stimulus luminance (varying from its minimum at L_{min} to its maximum at L_{max} about the mean luminance L_0), the dashed grey lines show the envelope that is being applied to the carrier grating. Dashed white lines in panels d) to f) show where the cross section is taken. Panels a) and d) show a full vertical Battenberg array with wavelength λ. Panels b) and e) show the array modulated to give the "black check" (no contrast in centre) condition with two cycles per check. Panels c) and f) show the array modulated to give the "white check" (contrast in centre) condition.

The array (shown in Figure 3.4d) is produced by multiplying this by the carrier grating

$$\mathbf{L}[x, y] = c \cdot \mathbf{E}[x, y] \cdot \sin\left[u(x - x_0) + v(y - y_0) + \phi\right]. \tag{3.31}$$

Elements from the array can be removed in a checkerboard pattern by a square-wave plaid modulator defined by the following equation

$$\begin{aligned}\mathbf{M}[x, y] = \mathrm{sgn}\langle\, &\cos[2\pi f_m(x\cos\theta_m + y\sin\theta_m) + \phi_m] + \ldots \\ &\cos[2\pi f_m(x\cos -\theta_m + y\sin -\theta_m) + \phi_m]\rangle,\end{aligned} \tag{3.32}$$

where f_m, θ_m and ϕ_m are the frequency, orientation and phase of the modulation. This is multiplied by the full Battenberg to produce the checked stimulus

$$\mathbf{L}[x, y] = c \cdot \mathbf{E}[x, y] \cdot \mathbf{M}[x, y] \cdot \sin\left[u(x - x_0) + v(y - y_0) + \phi\right]. \tag{3.33}$$

To produce stimuli made of arrangements of entire Battenberg elements, the values of f_m, θ_m and ϕ_m must be chosen so as to create a modulator that lines up with the zero-crossings in the

stimuli (see Figure 3.4, panels b-c and e-f).

For these stimuli, the effective signal area can be calculated as

$$\text{area}_{\text{stimulus}} = n_{\text{elements}} \times \text{area}_{\text{element}}. \tag{3.34}$$

3.5 Procedures

3.5.1 Two-interval forced-choice

Most of the experiments in this thesis (Chapters 4, 6 to 8) use a two-interval forced-choice (2IFC) procedure where each trials contain two temporal intervals, with each marked by a tone. In one randomly-chosen interval a stimulus is presented at a particular signal intensity (typically a contrast level here). The other interval contains a null stimulus which is absent of the feature that the observer is detecting, (e.g. blank for contrast detection). The observer's response is made by pressing one of two buttons to indicate which interval they think contains the target. Feedback on the correctness of the response can be provided by the pitch of an additional beep after the observer's response.

3.5.2 Single-interval identification

The orientation coherence experiment carried out in Chapter 9 uses a single-interval identification procedure. This is similar to the 2IFC case given above, however there is only one stimulus presentation interval and the observer's task is to identify which of two stimuli (or types of stimuli) are presented and respond accordingly.

3.5.3 Staircase methods

The behavioural response of observers to stimuli of different magnitudes, as sampled in psychophysical experiments, can be used to characterise a psychometric function. Collecting data for this purpose by using a process that tests an equal number of times at many stimulus levels lacks efficiency. It is often far quicker to use an adaptive method which favours testing at stimulus levels which will play a greater role in constraining the parameters of a fitted psychometric function (Kingdom & Prins, 2010). The simplest of these methods are staircase methods. These were first introduced by Dixon and Mood (1948) to study the sensitivity of explosives, but have since found wide use within many fields including psychophysics.

Staircase methods involve determining the next stimulus level to be tested by reacting to the results of previous trials. In the simplest case of a one-up one-down staircase, a correct response on the previous trial results in the next trial being tested with a lower stimulus magnitude (i.e. making the task more difficult) while an incorrect response results in the next trial being tested with a higher stimulus magnitude. As the staircase then overshoots the point at which the behaviour changes it will reverse direction again, eventually oscillating around a value of interest. For example, a one-up one-down staircase will tend to sample at stimulus magnitudes where the probability of a correct response is approximately 0.5 (staircases of this type are typically used for matching experiments where the point of subjective equality is the value to be determined). It possible therefore to use the average of the staircase reversals to calculate an estimate of the stimulus magnitude that corresponds to that P value, throughout this thesis however it is the parameters of a psychometric function fitted to the data which are reported (see Section 3.6 below).

Most of the studies in this thesis use three-down one-up staircases, which aim to converge at $P(\text{correct}) = 0.794$ (Wetherill, 1963; Wetherill & Levitt, 1965). Simulations have shown that the staircase may fail to reach this value (García-Pérez, 1998), however as the thresholds presented here are calculated using fitted psychometric functions the actual value that the staircases converge at is not critical.

For contrast detection experiments performed in Liberator, a pair of independent three-down one-up staircases were interleaved. The step sizes started at 12 dB and halved at each of the first two reversals to 6 dB and then to 3 dB. Response data were not recorded until after the second reversal. In Chapter 4, staircases terminated after 50 recorded trials. In Chapters 6 to 8 the staircases were limited to whichever was reached first of 70 trials or 12 reversals (the minimum number of trials with which 12 reversals can be reached is 48). Details for the staircases used in the orientation coherence study in Chapter 9 are provided in that chapter.

3.6 Psychometric functions

3.6.1 The psychometric function for detection

Psychophysical data which are expressed as the probability of the correct response to a stimulus (e.g. detection) as the function of changes in some relevant property of that stimulus (e.g. its contrast) are typically well-described by a curve that has a sigmoidal shape. There are several types of psychometric function with this shape that are used to fit psychophysical data. There are four degrees of freedom in these functions (shown in Figure 3.5). The lower asymptote is the guess rate, the expected performance level for an observer responding without using any information from the stimulus. The upper asymptote is the level where performance is limited only by observer "lapses", i.e. where they fail to perform the task correctly due to

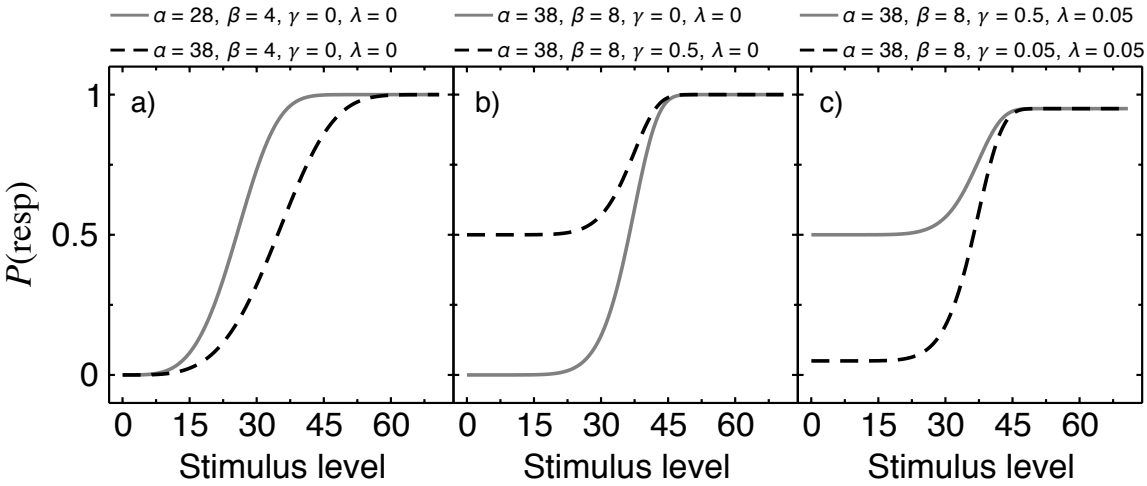

Figure 3.5: The effects of the parameters of a Weibull function. Starting with the original function (grey line) in panel a), the black dashed curve shows the effect of varying α, and panel b) shows the effects of varying β and γ. Panel c) shows the effect of varying λ and a function where $\gamma = \lambda$. Parameters for the curves are shown above the panels.

non-stimulus dependent factors such as inattention or blinking. The slope of the transition between the two asymptotes is how quickly the observer's behaviour changes from one performance level to the other based on the use of information from the stimulus. The location of the transition along the x-axis is controlled by the observer's sensitivity for performing this task (the stimulus level corresponding to a particular criterion performance level is given as the "threshold"). For the Weibull psychometric function these are controlled by the parameters γ, λ (the upper asymptote being at $1 - \lambda$), β and α respectively

$$\mathsf{f}(x) = \gamma + (1 - \gamma - \lambda) \cdot \langle 1 - \exp\left[-\left(\frac{x}{\alpha}\right)^{\beta}\right]\rangle. \tag{3.35}$$

Achieving reliable estimates of the relevant parameters is aided by fixing irrelevant parameters where possible. For 2IFC tasks, the guess rate (γ) can be fixed at 0.5, reducing the number of fitted parameters to three. For data collected using a staircase, the upper asymptote (controlled by λ) is not usually well-constrained and should be fixed (Prins, 2012). In the results from 2IFC experiments presented here λ is fixed at 0, making the assumption that the observer never lapses (would always respond correctly for a stimulus of sufficiently high contrast). Although this expectation is not realistic, lapse rates from experienced observers are expected to be very low, and the only effect of changing the fixed value of λ would be to globally bias the calculated slopes and thresholds (i.e. its effects are irrelevant when only comparisons between these parameters found for different conditions are made). This leaves the two free parameters controlling the threshold and slope (α and β), which are then found by fitting the psychometric function.

3.6.2 The psychometric function for identification

In identification experiments where the observer has to determine which of two classes the stimulus belongs to, psychometric functions plot the probability of identifying the stimulus as belonging to a particular class (e.g. horizontal as opposed to vertical) against an abscissa which ranges from a stimulus values inconsistent with that class (e.g. orientations away from the horizontal) to values consistent with that class (e.g. orientations approaching the horizontal). The psychometric function then describes the probability that the observer will respond to a stimulus with a particular value along that dimension by indicating that it is a member of one class or the other. The position of the psychometric function along the x-axis indicates the bias of the observer to respond by assigning the stimulus to a particular class, and the slope of the function is a measure of how sensitive the observer is at the identification task.

3.6.3 Probit

In Chapter 4, thresholds were derived from cumulative Gaussian psychometric functions fitted using the Probit method (Finney, 1971) in MATLAB. This fitting also gives the rejection criteria used in the contrast detection studies where data was averaged across repetitions (Chapters 4, 6 and 7). Where the standard error of the fit provided by this method to the data from a particular experimental session was greater than 3 dB the subject repeated the data collection for that session. This rejection criterion is a rule of thumb for avoiding the inclusion of data from sessions where unusually irregular behaviour was recorded. In the case where a condition that produced a rejected dataset was originally run interleaved with other conditions then all of those conditions were repeated.

3.6.4 Palamedes

In Chapters 6 to 9 psychometric function fitting was performed using the Palamedes toolbox (Prins & Kingdom, 2009; Kingdom & Prins, 2010). In Chapters 6 to 8, Weibull psychometric function fits were used for the sake of consistency with the analyses of results presented in previous studies. In Chapter 9, cumulative Gaussian psychometric functions were fitted to the orientation discrimination data.

3.7 Model fits

3.7.1 Fitting models to data

There are two different levels on which models of the visual system can be considered. These are the gross level where the architecture of the model describing its component stages and

how they are connected to each other is outlined, and the more detailed level where the parameters which control the precise behaviour of these components are specified. For example, including a filtering stage in a model is a matter of architecture, but the centre frequency and bandwidth of that filter would be a parameter. These model parameters are often varied to bring model predictions in line with empirical data, the purpose of which is either to find out what value of that parameter would be consistent with the data or to simply prevent the effects of that parameter (which may be irrelevant to the question at hand) from influencing the quality of the fit.

3.7.2 The downhill simplex method

Models are fitted to data in this thesis with the downhill simplex method (Nelder & Mead, 1965; Press, Flannery, Teukolsky, & Vetterling, 1989), implemented using the *fminsearch* function in MATLAB. The purpose of this method is to vary the model parameters and find the set which results in the lowest RMS error between the model and the data. Briefly, the method creates a simplex of $n + 1$ points on an n-dimensional surface, where the elevation at each location is the RMS error between the data and the model prediction with that location's particular combination of values for the n parameters. The simplex then "walks" downhill, primarily by replacing the point in the simplex with the highest elevation (RMSe) with its reflection on the other side of the simplex. Using this method, minima in the surface can be found. For situations where multiple minima are likely, the chance of finding a global minimum is increased by jittering the location of the simplex when it is stuck in a local minimum, and by running the routine several times with different starting points.

CHAPTER 4

The visual field inhomogeneity in contrast sensitivity

4.1 Motivation and summary

Investigating the summation of threshold contrast across the visual field requires an understanding of how contrast sensitivity varies as a function of eccentricity. Sensitivity has previously been reported to decline as a linear function of eccentricity when plotted on log-lin axes. Here, this account is challenged following experiments performed with a higher sampling density in the central nine degrees of the visual field. The findings instead suggest a two-part bilinear relationship composed of an initial steep slope followed by a shallower decline. A method for producing an attenuation surface of this shape is put forward, along with a set of parameters derived from the experimental results.

4.2 Introduction

4.2.1 The visual field inhomogeneity in contrast sensitivity

Visual perception begins with the transduction of patterns of luminance incident on the retina by photoreceptors. The neural architecture is not homogenous however, with the distribution of photoreceptors and the degree of convergence in their outputs varying over the retina (Perry & Cowey, 1985; Curcio & Allen, 1990; Curcio et al., 1990). This, combined with other inhomogeneities in the processing after the retina (Daniel & Whitteridge, 1961; Rovamo, Virsu, & Näsänen, 1978; Rovamo & Virsu, 1979; Virsu & Rovamo, 1979), is thought to lead to better performance in the centre of the visual field than in the periphery for most tasks (see Strasburger et al., 2011, for an extensive review). Here, the intent is to measure the variation in

contrast sensitivity across the visual field in order to create a surface that describes the inhomogeneity. Such a surface might then be used in computational models of the visual system (see Chapter 5), or to transform stimuli before presenting them to observers in order to compensate for the effects of the inhomogeneity (see Chapters 6 and 7).

The effect of eccentricity on contrast sensitivity has been investigated extensively in the past (Pöppel & Harvey, 1973; Hilz & Cavonius, 1974; Koenderink, Bouman, Bueno de Mesquita, & Slappendel, 1978a,b,c,d; Rovamo et al., 1978; Rovamo & Virsu, 1979; Rijsdijk, Kroon, & van der Wildt, 1980; Robson & Graham, 1981; Wright & Johnston, 1983; Kelly, 1984; Johnston, 1987; Pointer & Hess, 1989; Rovamo, Franssila, & Näsänen, 1992; Foley et al., 2007; Hess, Baker, May, & Wang, 2008). A summary of the methodology used in the most relevant previous literature is provided in Table 4.1. Previous studies typically report a linear decline in log contrast sensitivity with increasing eccentricity (distance from the point of fixation). For those studies that also tested spatial frequency effects (e.g. Robson & Graham, 1981; Pointer & Hess, 1989), contrast sensitivity declined with eccentricity at roughly the same rate when the eccentricity was expressed relative to the period of the stimulus spatial frequency i.e. in "carrier cycles" (as opposed to being expressed in degrees of visual angle).

Where previous studies have tested contrast sensitivity over several meridians, they have found the decline in sensitivity to be anisotropic. Pointer and Hess (1989) found that sensitivity declined more steeply along the vertical meridian than along the horizontal meridian. There have also been several reports of a superior-inferior anisotropy, in which performance is better at isoeccentric locations in the inferior hemi-meridian than in the superior hemi-meridian (reviewed in Abrams, Nizam, & Carrasco, 2012). Any description of how contrast sensitivity varies over the visual field must therefore take account of these anisotropies.

4.2.2 Effects of stimulus orientation

There have been several previous studies reporting that the visual system has different sensitivities to stimuli of different orientations. These stimulus orientation effects come in two different forms, where they are based either on the absolute orientation of the stimulus or on the orientation of the stimulus relative to the polar angle of its location in the visual field. The possibility of both absolute and relative orientation effects occurring in the region of interest is investigated here in order to determine whether it is necessary to include these effects in a model of the central visual field's contrast sensitivity to oriented stimuli.

The absolute orientation effect reported previously is the "oblique effect", where horizontally- and vertically-oriented stimuli are more easily detected than oblique stimuli at medium to high spatial frequencies (Campbell, Kulikowski, & Levinson, 1966; Berkley, Kitterle, & Watkins, 1975; Heeley & Timney, 1988; Long & Tuck, 1991). A relative orientation effect was first reported by Rovamo, Virsu, Laurinen, and Hyvärinen (1982), who found a "meridional resolution

Publication	SF (c/deg)	Stimulus type / window	Size	# samples within 4.5 deg	Orientation	Eccentricity	Meridian
Koenderink et al. (1978a)	2 - 25.5	Square	0.5 x 0.5 deg	4	Horizontal	0 - 80 cycles 0 - 8 deg	H
Rovamo et al. (1978)	1 - 32	Semicircle	Radius = 1 deg	3	Vertical	0 - 60 cycles 0 - 30 deg	H & V
Rijsdijk et al. (1980)	0.35 - 6	Plaid patch	1 x 1 cycles	4	H x V	0 - 36 cycles 0 - 6 deg	H, V & D
Robson and Graham (1981)	1.5 - 24	Raised-cosine	4 x 8 cycles	3	Horizontal	0 - 32 cycles 0 - 21 deg	V
Wright and Johnston (1983)	0.25 - 9	Rectangle	3.5 x 0.7 degrees	4	Vertical	0 - 72 cycles 0 - 12 deg	V
Kelly (1984)	0.5 - 16	Annulus	12 - 500 deg^2	2	Concentric	0 - 12 deg 0 - 36 cy	N/A
Johnston (1987)	2 - 12	Square	12 x 12 cycles	2	H & V	0 - 40 deg 0 - 480 cycles	H
Pointer and Hess (1989)	0.05 - 12.8	Gabor	σ = 3.2 cycles	1	Horizontal	0 - 96 cycles 0 - 60 deg	H & V
Rovamo et al. (1992)	3	Not specified	6 x 6 cycles	2	Vertical	0 - 39 cycles 0 - 13 deg	H
Foley et al. (2007)	4	Gabor	σ = 1 cycle	4	Vertical	0 - 20 cycles 0 - 5 deg	H
Hess et al. (2008)	0.5 - 3	Gabor	σ = 1 cycle	1	Horizontal	0 - 30 cycles 0 - 60 deg	H
This study	0.7 - 4	Log-Gabor	sfbw: 1.6 oct oribw: +/- 25 deg	4 - 9	H, V & D	0 - 18 cycles 0 - 26 deg	H, V & D

Table 4.1: Summary of stimulus and experiment details from previous studies that investigated the decline in contrast sensitivity with eccentricity. Where studies consisted of multiple experiments the details shown are for the experiments most similar to those conducted in this study. Abbreviations: H = Horizontal, V = Vertical and D = Diagonal.

effect" which was dependent on the stimulus orientation relative to the angle at which it appeared from the point of fixation. Subjects had higher acuity for grating patches which were radially-aligned (where the bars of the grating were parallel to a line drawn from fixation to the grating patch) and lower acuity for patches orthogonal to those. This effect became more pronounced with increasing distance from fixation (strong effects were found beyond 20 degrees of eccentricity). This relative orientation effect on sensitivity has also been found in other stimulus modalities. Pointer and Hess (1989) did not find a higher contrast sensitivity for patches aligned in this way, however such an effect has been reported by Sasaki, Rajimehr, Kim, Ekstrom, Vanduffel, and Tootell (2006). Due to these contradictory previous results, this issue is investigated here.

4.2.3 This study

The results from this study show that the decline in log contrast sensitivity in the central visual field is non-linear and is best fit by a bilinear function, where the initial slope is approximately twice as steep as the subsequent shallower decline. The scale invariance found by previous investigators (e.g. Pointer & Hess, 1989) is confirmed, as are the two visual field anisotropies (both the horizontal-vertical and superior-inferior anisotropies from Abrams et al., 2012). The results show that the stimulus orientation effects, if present at all, are small enough to not be a necessary feature in a model of the central nine degrees of the visual field. This allows a simple two-dimensional attenuation surface to be developed via radial interpolation of the bilinear functions that are fitted to the data, providing a generalised map of contrast sensitivity to stimuli of arbitrary orientation in the central visual field.

4.3 Methods

4.3.1 Equipment

Three experimental set-ups were used. In each case, stimuli were stored in a CRS ViSaGe and presented on a gamma-corrected CRT monitor (Nokia Multigraph 445X, Philips MGD403, or Eizo Flexscan T68). All monitors had a refresh rate of 120 Hz, and mean luminances varied from 60 to 85 cd/m^2 between the monitors. The stimuli had 12 pixels per carrier cycle for spatial frequencies of 2 to 4 c/deg (sufficient to avoid luminance artefact problems that might arise from adjacent pixel non-linearity, see García-Pérez & Peli, 2001). The viewing distance for the 4 c/deg stimuli was 1.19 metres. At this distance, 48 pixels on the screen subtended 1 degree of visual angle. The viewing distance was adjusted to scale the retinal image to the desired spatial frequency (59.5 - 119 cm for the range 2 - 4 c/deg). For stimuli with a spatial

frequency below 2 c/deg, the stimulus was first doubled in size on the screen (24 pixels per carrier cycle) and the viewing distance was adjusted appropriately (41.7 - 83.3 cm for the range 0.7 - 1.4 c/deg).

In Experiment 1, the two principal observers (ASB and DHB) used different equipment setups. To ensure that this was not responsible for any differences in their results both observers ran a subset of the Experiment 1 conditions on the other equipment to that on which they collected their Experiment 1 data. The results from each laboratory were found to be in agreement within observer, rather than being dependent on which equipment was used.

4.3.2 Stimuli

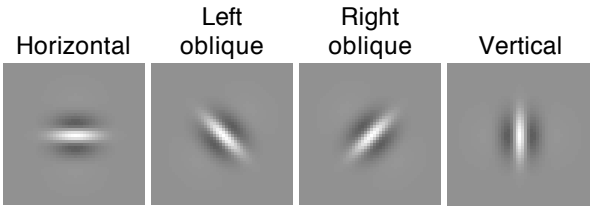

Figure 4.1: Cartesian-separable log-Gabor stimuli generated in cosine-phase with orientations of (left to right): 90°, 135°, 45° and 0°.

The stimuli for this experiment were luminance-modulated cosine-phase log-Gabors patches (see General Methods chapter, Section 3.4). For the main experiments here, the stimuli used had a spatial frequency bandwidth of 1.6 octaves (full width at half height) and an orientation bandwidth of $\pm 25°$ (half widths at half height). These bandwidths were chosen in order to attempt to match the stimuli to the receptive fields found in V1 (see Section 3.4.5). The log-Gabors used in this study were presented at four different orientations (horizontal: 90°, left oblique: 135°, right oblique: 45° and vertical: 0°; see Figure 4.1) and six spatial frequencies (0.7, 1, 1.4, 2, 2.8 and 4 c/deg). The stimulus duration was 100 ms. Stimulus contrasts were calculated as delta-contrast and expressed in dB re 1% (see Section 3.2).

4.3.3 Observers

Data were collected from four observers: ASB, DHB, SAW and TSM. The observers were 22, 28, 44 and 46 years old respectively, and wore optical correction appropriate for the viewing distances tested when required. Experiments were performed binocularly with natural pupils.

4.3.4 Procedures

In Experiment 1, 4 c/deg log-Gabor stimuli of all four orientations were presented at four eccentricities (0, 6, 12 and 18 cycles; or 0, 1.5, 3 and 4.5 degrees of visual angle) along eight hemi-

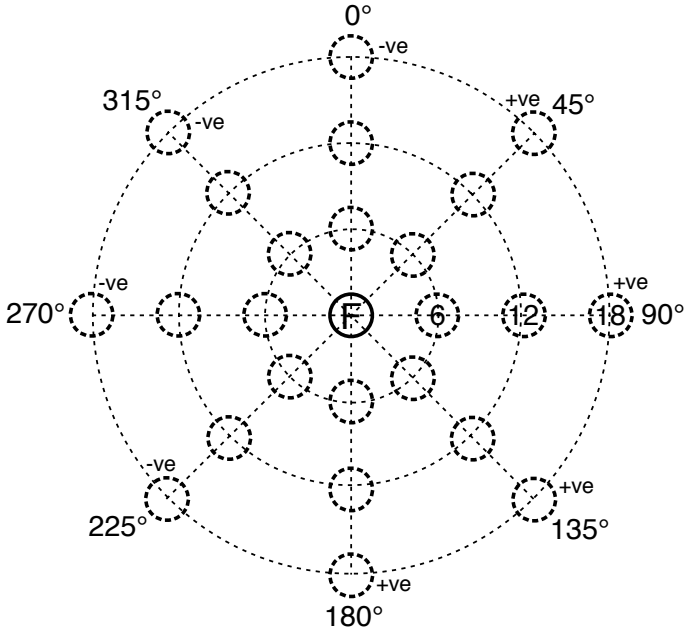

Figure 4.2: Diagram showing the four meridians tested in these experiments, and the eccentricities (in carrier cycles) along those meridians that were used in Experiment 1. "F" marks the fixation circle. "+ve" and "-ve" labels refer to the direction along the meridian that is plotted in the graphs below.

meridians (0°, 45°, 90°, 135°, 180°, 225°, 270° and 315°) radiating from the centre of the visual field. A diagram of these locations is shown in Figure 4.2. Stimuli were blocked such that in each session thresholds were only being determined for a stimulus of a single orientation at a single position in the visual field (i.e. this design does not feature any extrinsic uncertainty). This gave 100 blocks (the 4 patch orientations in Figure 4.1 at each of the 25 locations in Figure 4.2), which were all repeated in a randomised order four times by observers ASB and DHB. Two more observers (SAW and TSM) provided additional data for a subset of the conditions.

Thresholds were measured using a two-interval forced-choice (2IFC) three-down, one-up staircase procedure (see General Methods chapter, Section 3.5). Feedback was provided. Each condition was repeated four times by each observer. Contrast detection thresholds for each repetition were calculated using a probit fit to the staircase data (see General Methods, Section 3.6), allowing the mean and standard error to be calculated across repetitions.

To reduce extrinsic uncertainty, there was a continuously visible circle (diameter of 3 carrier cycles, line width of 1 pixel, with a contrast of 25%) placed to surround the location where the target would appear. An identical circle was also used for fixation, such that in the condition where the target was at fixation the two circles were coincident. In Experiment 2, the circles were replaced with pairs of dots as described in the results section for that experiment.

For the majority of the experiments the observer fixated in the centre of the monitor, with the stimuli appearing at a location on the screen relative to this central fixation. This differs from

the more conventional method where the stimuli are located in the centre of the display in order to avoid potential problems that might arise from inhomogeneities in the monitor (such as variations in luminance across the display, see García-Pérez & Peli, 2001). A subset of conditions from Experiment 1 were retested where the target appeared in the centre of the screen and the fixation point was placed at various positions on the monitor to control where in the visual field the stimulus was presented. The results from the two different experiments were generally in agreement, with no systematic differences between them.

The methods for the other experiments were similar to those for Experiment 1, with the differences described in the relevant part of the Results section.

4.4 Results from Experiment 1

4.4.1 Contrast sensitivity across the central visual field

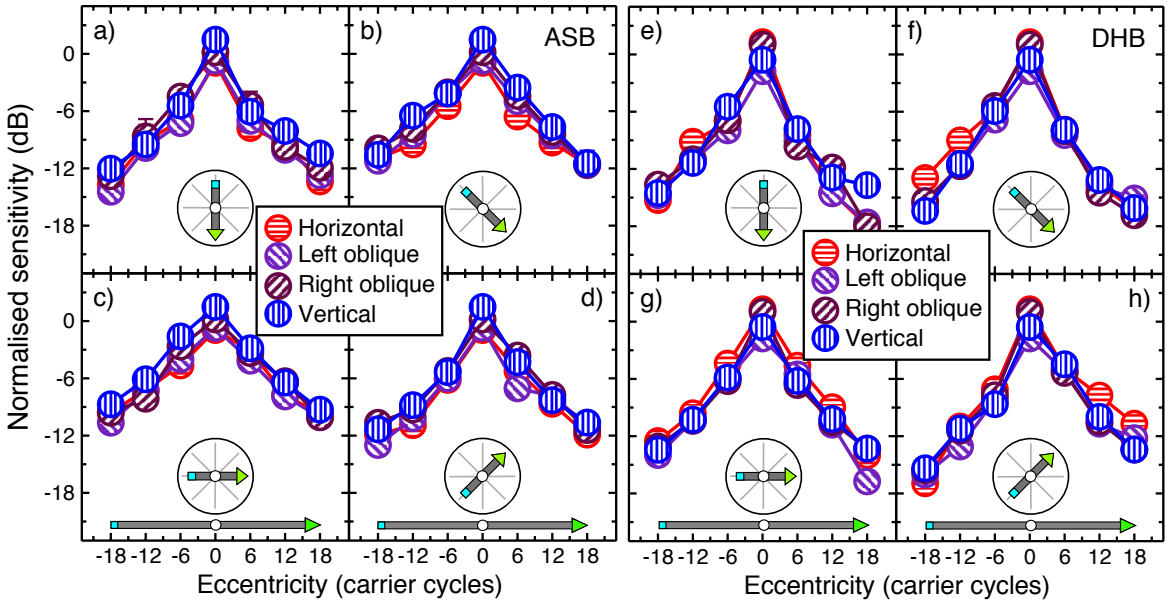

Figure 4.3: Contrast sensitivity data from Experiment 1 for observers ASB (left) and DHB (right), normalised to the observer's sensitivity at fixation. The separate plots show the sensitivity data for the detection of a 4 c/deg log-Gabor patch for each meridian, as indicated by the diagram in each plot. The four sets of symbols show data from the different stimulus orientations (as indicated in the legend). Eccentricity is expressed in stimulus carrier cycles, the visual angle of the range shown here is +4.5 degrees to -4.5 degrees along each meridian. Error bars in this figure (and in all subsequent figures) show ± 1 standard error where visible. Where they are not visible this is due to their being smaller than the symbol size.

The contrast sensitivity along the eight tested hemi-meridians for two observers can be seen in Figure 4.3. In agreement with previous findings, the sensitivity is greatest at fixation, and

	Individual hemi-meridian fall-off fits (dB/cycle)							
Angle	0°	180°	90°	270°	45°	135°	225°	315°
ASB	0.72	0.65	0.55	0.55	0.61	0.64	0.65	0.59
DHB	0.79	0.92	0.79	0.74	0.70	0.90	0.86	0.84
Average	Vertical		Horizontal		Diagonal			
ASB	0.69 ±0.04		0.55 ±0.00		0.62 ±0.01			
DHB	0.86 ±0.07		0.77 ±0.03		0.83 ±0.04			

Table 4.2: Log-contrast sensitivity decline gradient (in dB) for linear fits to Experiment 1 data. Averages are given as the mean ±1 standard error.

the decline is steeper along the vertical meridian than along the horizontal meridian. The diagonal meridians appear to show declines in sensitivity that are intermediate between those for the horizontal and vertical meridians (a point which will be returned to in the modelling). The overall declines were slightly steeper than those found in previous studies, the gradients from a linear fit ($y = mx + c$, with both m and c allowed to vary freely) are shown in Table 4.2.

To test for the absolute and relative orientation effects, a pair of 3-way repeated-measures ANOVA tests was performed using PASW Statistics (version 18.0, IBM) for each observer. The first test was for factors of: eccentricity, hemi-meridian, and absolute patch orientation. The second test was the same, but with patch orientation defined relative to the meridian that the stimulus was placed on. For example, a right oblique patch placed on the 225° to 45° meridian (see Figure 4.3) would have an absolute orientation of 45°, but a relative orientation of 0°. Results from Mauchly's test of sphericity showed that the ANOVAs that found significant results did not suffer from a violation of sphericity, therefore no correction was required.

4.4.2 Absolute orientation effects

Significant effects of absolute patch orientation were found for both ASB (Mauchly's test n.s. $\chi^2(5) = 5.45$, $p = 0.43$; ANOVA $F_{3,9} = 11.50$, $p < 0.01$) and DHB (Mauchly's test n.s. $\chi^2(5) = 4.01$, $p = 0.61$; ANOVA $F_{3,9} = 5.52$, $p = 0.02$), but were small and not consistent across observers. ASB was most sensitive to vertical patches and least sensitive to horizontal patches (compare blue and red symbols in panels a-d of Figure 4.3), whereas DHB was most sensitive to horizontal patches and least sensitive to left-oblique patches (compare red and magenta symbols in the panels e-h of Figure 4.3).

In addition to the ANOVA, the absolute orientation effects were further investigated by three paired Bonferroni-corrected t-tests per observer. In the first two tests the thresholds for orthogonal patch orientations were compared to each other (horizontal against vertical, and left-oblique against right-oblique), paired by eccentricity and hemi-meridian (i.e. comparing the thresholds for orthogonal patches at the same location in the visual field). For these analyses, ASB showed significant differences in sensitivity for both comparisons ($p < 0.01$), whereas

DHB showed no significant difference in either the horizontal vs. vertical (p = 0.24) or the left- vs. right-oblique (p = 0.09) tests. In the third test the results for the two cardinal stimulus orientations (horizontal and vertical) were compared against those for the two oblique orientations by averaging the pairs of thresholds from each repetition (i.e. a t-test comparing the average of the horizontal and vertical thresholds against the average of the two oblique thresholds at each of the tested locations in the visual field). For ASB there was a small difference (0.13 dB) between the thresholds of the cardinal and oblique patches, however this was not significant (p = 0.23). For DHB the oblique effect was larger (0.66 dB) and achieved significance (p < 0.01).

The oblique effect reported here is smaller than that found previously at higher spatial frequencies (e.g. Campbell et al., 1966, found effects of 2 - 6 dB for spatial frequencies in the range of 10 - 30 c/deg) but consistent with that found at similar spatial frequencies to those tested here (see also Long & Tuck, 1991). Based on the findings here it is concluded that for these stimulus conditions the absolute orientation effects were of little concern, since when they were statistically significant they were inconsistent across observers and small in size.

4.4.3 Relative orientation effects

Comparing the mean contrast detection threshold of patches aligned with the meridian they were placed on with that of patches having the orthogonal orientation, I found a small radial advantage for both ASB (0.25 dB) and DHB (0.56 dB). The ANOVA that tested relative orientation (across all four orientations) found this effect to be non-significant for ASB (ANOVA $F_{3,9}$ = 1.54, p = 0.27), but significant for DHB (Mauchly's test n.s. χ^2(5) = 0.44, p > 0.99; ANOVA $F_{3,9}$ = 4.07, p = 0.04). However, despite the overall significant result on the ANOVA for DHB there were no significant pairwise comparisons across aligned and orthogonal patches.

Previous studies that have reported an effect of relative orientation (e.g. Rovamo et al., 1982; Sasaki et al., 2006) investigated greater eccentricities (25 deg and 15.5 deg, respectively) than those tested here, so it is possible that relative orientation effects might become stronger with increasing eccentricity. If this were the case within the tested eccentricity range for Experiment 1, then it would appear in the two-way interaction between eccentricity and relative orientation. The results of this analysis however did not show a significant effect for either observer (ASB: ANOVA $F_{6,18}$ = 0.59, p = 0.74; DHB: ANOVA $F_{6,18}$ = 2.09, p = 0.11).

4.4.4 Results for vertical stimuli

The lack of substantial and consistent differences in sensitivity to patches of different orientations means that the contrast sensitivity findings for just one patch orientation can be generalised across all patch orientations within the region of interest. Two additional observers

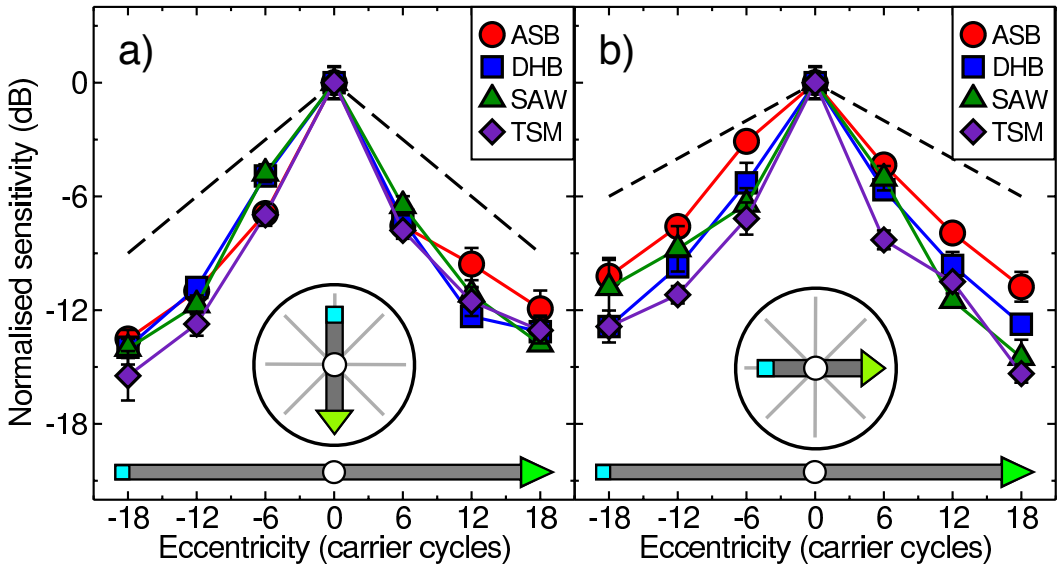

Figure 4.4: Contrast sensitivity across the cardinal meridians for all four observers. Vertical log-Gabor stimuli with a spatial frequency of 4 c/deg were used. The black dashed lines presented here for comparison are the gradients for the vertical (0.5 dB/cycle) and horizontal (0.33 dB/cycle) meridians reported by Pointer and Hess (1989).

(SAW and TSM) performed a subset of conditions from Experiment 1. Their data, for vertical patches in the horizontal and vertical meridians, can be seen in Figure 4.4 along with the data from ASB and DHB for those conditions (replotted from Figure 4.3). The sensitivity decline gradients reported previously by Pointer and Hess (1989) are shown by the dashed lines in Figure 4.4. A comparison of these lines with the data from this study shows that the rate of decline found here is considerably steeper than that reported by Pointer and Hess (1989).

In Figures 4.3 and 4.4 it appears that the decline in log contrast sensitivity with distance from fixation is non-linear. For most observers and hemi-meridians there is a steep initial decline followed by a shallower slope. However the sampling for this dataset is not sufficiently fine to either support this interpretation over other nonlinear declines, or to pin down the location of the transition between these two slopes (the "knee-point", see the Modelling section) beyond placing it within the first 18 cycles (4.5 degrees). The shape of the decline, the location of the knee-point, and whether the position of the transition is based on periods of the stimulus or absolute position on the retina will be investigated in the results from Experiments 2 and 3.

4.5 Results from Experiment 2

4.5.1 Finely-spaced mapping of contrast sensitivity

Experiment 2 was conducted in order to determine the location of the transition between the two stages of the bilinear decline in log contrast sensitivity. A subset of the conditions of Ex-

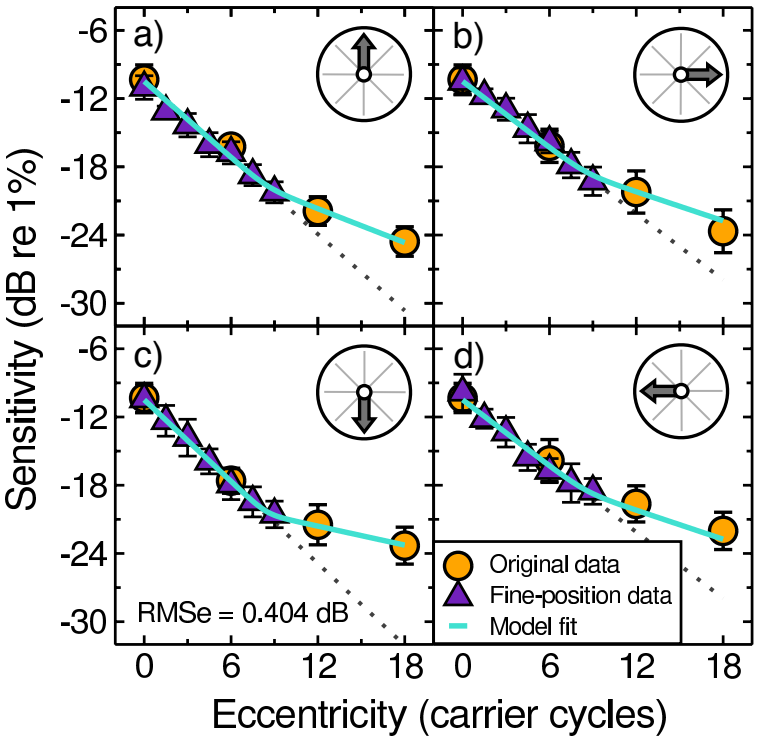

Figure 4.5: Fine-positioning data from Experiment 2 (triangles), plotted with the data from Experiment 1 (circles). The data presented here are averaged across the four observers (ASB, DHB, SAW and TSM), and the four panels are fitted simultaneously with the eight-parameter witch's hat bilinear model (see Modelling section). Parameters for the model fit shown here are provided in Table 4.4. The dotted lines extrapolate the initial (m_1) decline.

periment 1 were repeated with finer sampling. A vertical stimulus orientation was used on just the horizontal and vertical meridians and the eccentricities were more closely spaced (0 - 9 cycles in intervals of 1.5 cycles). The fixation and target location circles were replaced by two flanking pairs of dots, as continuing the use of circles would result in them overlapping in the conditions where the target was near fixation. This also required the pair of dots to be oriented horizontally when the vertical meridian was being tested, and vertically in the conditions that tested the horizontal meridian. This experiment was performed by all four observers.

The results from Experiment 2 are shown in Figure 4.5 (triangles), along with a subset of the data from Experiment 1 (circles). The model fits will be addressed in a later section. The data shown here are averaged across all four observers. There is a strong agreement between the data from this experiment and that from Experiment 1, despite the differences between the two in terms of how fixation was managed. Figure 4.5 shows that the decline in contrast sensitivity is linear out to approximately 9 cycles, and after this point becomes more shallow (compare the data to the dotted lines which extrapolate the initial decline along each hemi-meridian). The location of this knee-point is investigated in the modelling section below.

4.6 Results from Experiment 3

4.6.1 Effects of spatial frequency

Experiments 1 and 2 provide evidence that the decline in log contrast sensitivity is bilinear, with the transition between the initial sleep and subsequent shallow declines occurring at a knee-point. As these two experiments only used stimuli of one spatial frequency however it is not possible to determine from their results whether the location of this knee point is absolute, or whether it depends on the scale of the stimulus. For the first case, the knee point might have a fixed location in degrees of visual angle (possibly based on retinal anatomy, see Section 4.7.6), for the second case it may be a fixed number of stimulus carrier cycles ("scale invariant").

Experiment 3 was designed to investigate the relationship between the shape of the decline in sensitivity and the spatial frequency of the target. It was conducted based on a subset of the conditions from Experiment 1, using horizontal log-Gabor patches of six spatial frequencies (0.7, 1, 1.4, 2, 2.8 and 4 c/deg) surrounded by continually visible circles to indicate their position. The 180° (inferior) hemimeridian was sampled at twice the frequency as in Experiment 1 (i.e. at 0 to 18 cycles in intervals of 3 cycles).

The results from Experiment 3 are shown in Figure 4.6 (the same data are plotted in the two columns against different x-axes). In agreement with Pointer and Hess (1989) the slopes appear parallel when eccentricity is expressed in stimulus carrier cycles (left column) but diverge when expressed in degrees of visual angle (right column), and the observers are less sensitive to the higher spatial frequencies than to the lower spatial frequencies. It appears from Figure 4.6 that the knee point has an eccentricity which is fixed in terms of being a constant number of carrier cycles from fixation. This, and the fitting of the model curves, is considered more closely in the Modelling section below.

4.7 Modelling

4.7.1 Bilinear model equations

The data shown in Experiments 1 to 3 appear to have a decline in log sensitivity which is bilinear. A function of this shape can be defined as

$$S = -\log_{10}\left(\frac{10^{m_1 E}}{10^{(m_1-m_2)\nu} + 10^{(m_1-m_2)E}}\right) + K, \tag{4.1}$$

where S is the contrast sensitivity (expressed in dB) and E is the eccentricity expressed in either degrees of visual angle or in stimulus carrier cycles (e.g. compare the two columns in

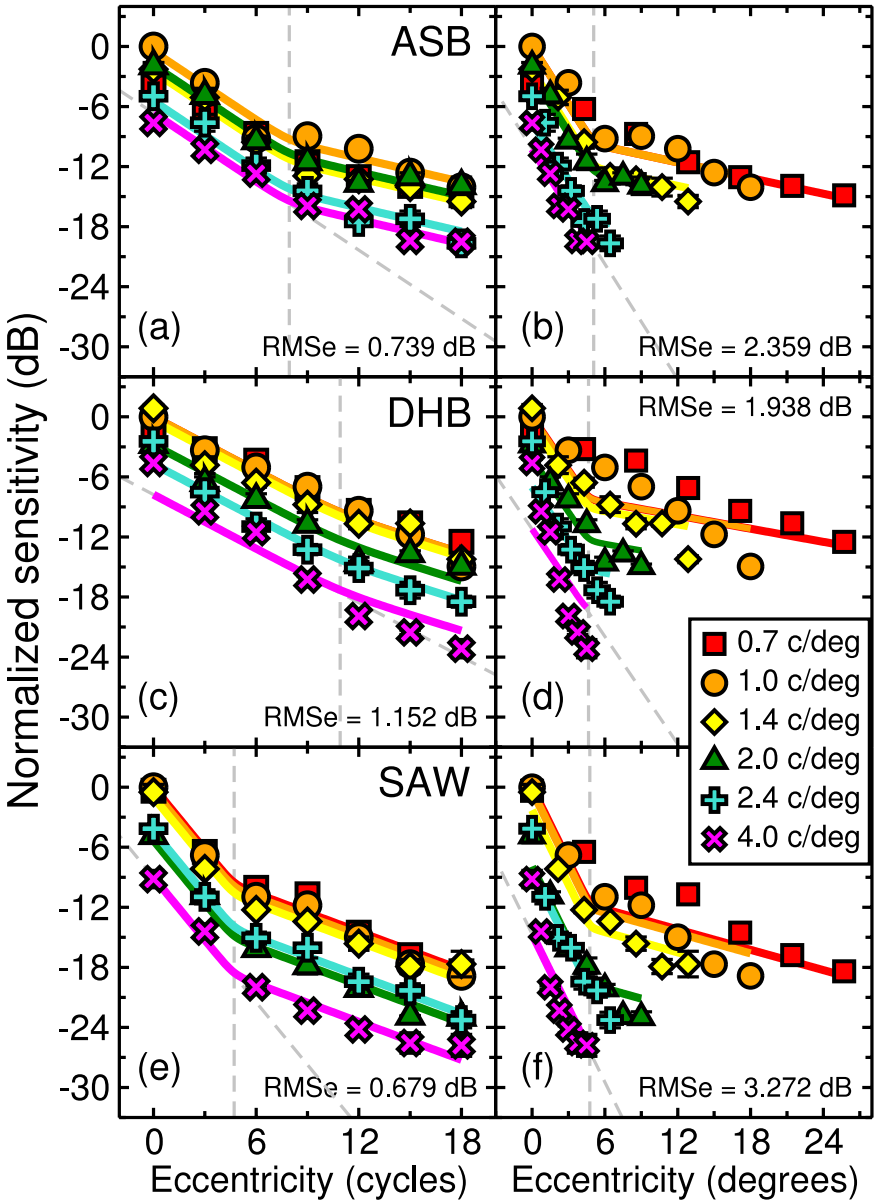

Figure 4.6: Contrast sensitivity decline data from Experiment 3 for six different spatial frequencies (indicated by their different symbols in the legend). Data are presented for the observers three ASB, DHB and SAW in separate rows. The two columns show the same contrast sensitivity data plotted against eccentricity expressed in stimulus carrier cycles on the left (panels a, c and e), and in degrees of visual angle on the right (panels b, d and f). The solid curves are witch's hat bilinear model fits to the data (see Modelling section). The parameters for these fits are provided in Table 4.6. The grey dashed lines show the positions of the fitted knee points (ν) and extrapolate the initial decline (with a negative gradient of m_1).

Figure 4.6). The free parameters are m_1 and m_2 which describe the slopes of the first and the second limbs of the bilinear function respectively (they are the gradients of the negative slopes, in dB per unit of eccentricity), ν which controls the location of the knee point (in the same units that E is expressed in) and K which controls the vertical offset of the entire function. For convenience I transfer the degree of freedom from K to k_2, where $K = k_1 + k_2$ and

$$k_1 = \log_{10}\left(\frac{1}{10^{(m_1-m_2)\nu}+1}\right). \tag{4.2}$$

Using k_2 as the vertical offset parameter in this way is useful because that free parameter is then the sensitivity of the observer to a stimulus presented at fixation, expressed in dB re 1%.

A peculiar property of Equation 4.1 is that the meanings of the m_1 and m_2 values switch around depending on which of the two is larger. The larger value will control the slope of the initial steeper slope, and the smaller will control the slope of the subsequent decline. This can be demonstrated by differentiating the function, revealing that

$$\frac{dS}{dE} = -\frac{m_1 10^{m_1\nu+m_2E} + m_2 10^{m_2\nu+m_1E}}{10^{m_1\nu+m_2E} + 10^{m_2\nu+m_1E}}. \tag{4.3}$$

A function that allowed the meaning of the parameters to switch around in this way could cause problems in fitting the models to the data. For this reason, the values of m_1 and m_2 are constrained such that $m_1 \geq m_2$.

It was desirable to compare the "concave" bilinear function (Equation 4.1) against a "convex" bilinear function (i.e. one that flexed in the opposite direction, with an initial shallower slope followed by a steeper slope). Equations 4.1 and 4.2 can be modified to produce such a function where

$$S = \log_{10}\left(\frac{10^{-m_1E}}{10^{-(m_1-m_2)\nu} + 10^{-(m_1-m_2)E}}\right) + K, \tag{4.4}$$

and

$$k_1 = -\log_{10}\left(\frac{1}{10^{-(m_1-m_2)\nu}+1}\right). \tag{4.5}$$

For Equations 4.4 and 4.5, m_1 and m_2 are constrained such that $m_1 \leq m_2$. This is necessary to simplify the fitting for the same reasons as given above for the convex bilinear function. Radial interpolation of the convex (Equations 4.1 and 4.2) and concave (Equations 4.4 and 4.5) bilinear functions produces surfaces that have the three-dimensional shape of a witch's hat (shown in Figure 4.14) and a Samurai hat respectively.

4.7.2 Bilinear model comparison

The data from the three main experiments and the two control experiments ("target central on monitor" and "equipment swap") were fitted by the two different bilinear functions in MATLAB

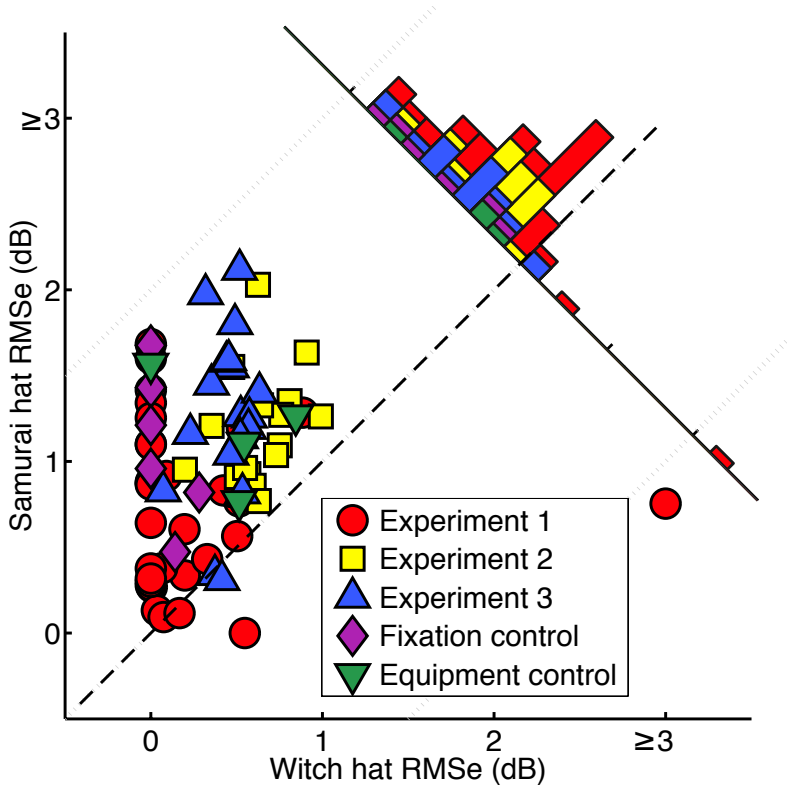

Figure 4.7: Comparison of the quality of the fits to the data provided by using the concave witch hat vs. the convex Samurai hat bilinear functions. Data are shown from the three main experiments and the two control experiments, with each hemi-meridian from each dataset being fitted independently by both models (each with four free parameters). The RMS errors from the two fits are plotted against each other in this scatter plot. For the purposes of presentation here, RMS errors larger than 3 dB were set to 3 dB (altering the position of a single data point so that it appears on the graph). The histogram shows the distribution of the differences between the witch's hat and Samurai hat model RMS errors, using the same colour code as the scatter plot.

using the downhill simplex method (see Section 3.7). The RMS errors from these fits were then compared in order to evaluate the goodness of the fit for each model. The data supplied to the fitting algorithm were from a single patch orientation on a single hemi-meridian in each case, and they were fitted independently with the four free parameters: m_1, m_2, ν and k_2.

The comparison of the RMS errors is shown in Figure 4.7. If the decline in sensitivity were linear, then it would be expected that the fits from the two bilinear functions in each case would be identical (because for a linear fit $m_1 = m_2$), as would the RMS errors. This would result in the points on the graph falling on the line $y = x$. For those cases where one fit is superior to the other the points will either lie to the left of that line (when the fit from the concave "witch's hat" fit is superior) or to the right of that line (if the fit from the convex "Samurai hat" fit is superior). In most cases (71 out of 76) the data were better fit by the witch's hat function than by the Samurai hat function. Performing a sign test for each experiment showed that this advantage was highly significant for the three main experiments ($p < 0.001$), and significant for

the control experiment where sufficient data had been gather for it to be so ($p = 0.031$ for the fixation control, $p = 0.125$ for the equipment control).

4.7.3 Determining the number of necessary model parameters

Model description (all include global k_2 parameter)	Free param.	RMSe (dB)	AIC
Separate m_1, m_2 and ν for each direction.	13	0.38	-13.4
Separate vertical m_1, m_2 and ν, combined for horizontals.	10	0.40	-17.7
Combined m_1, m_2 and ν for horizontals and for verticals.	7	0.45	-19.0
Global m_1, m_2 and ν.	4	0.69	-7.5
Combined m_1 and m_2 for horizontals, global ν.	8	0.40	-21.2

Table 4.3: Number of parameters, RMS error, and Akaike's Information Criterion for five different versions of the witch's hat bilinear model which were used to fit the combined data from Experiments 1 and 2 (data were averaged across the four observers). The variant in the bottom row is the preferred model (it has the lowest AIC score), its parameters are given in Table 4.4 and it is shown fitted to the data in Figure 4.5.

The bilinear function (Equation 4.1) provides a good fit to data from individual meridians (Figure 4.7), it is likely however that there are some model parameters that can be generalised across meridians. The most obvious example of this is the k_2 parameter, which must by definition be the same for all meridians (as it is the sensitivity at fixation). Other candidates for parameters which might be collapsed across meridians would be those that might be mirrored (i.e. the left and right hemi-meridians could have the same sensitivity declines). Reducing the number of parameters in this way has a practical benefit, as it allows a surface describing the attenuation across the central visual field to be derived from fewer data points.

Five models investigating potential symmetries were designed and are listed in Table 4.3, each having different numbers of free parameters. Allowing the parameters (excluding k_2) to vary freely in every direction produced the most irregular surface (13 parameters), whereas fixing the decline to be the same in every direction produced a surface which was rotationally symmetric (4 parameters). These models were fitted to the data from Experiments 1 and 2 (averaged across observers) using a simplex algorithm, and the RMS errors were compared in order to judge the models against each other. The assessments of the model were based not on which model fit closest to the data, but on which was the most parsimonious (i.e. balancing the quality of the fit provided against the number of free parameters). This was assessed using Akaike's Information Criterion (AIC; Akaike, 1974; Peirce, 2007) which is defined as

$$\text{AIC} = n \log(RMSe) + 2p, \tag{4.6}$$

where n is the number of data points, $RMSe$ is the error of the model fit to the data, and p is the number of model parameters. The preferred model is the one with the lowest AIC score.

From Table 4.3 it can be seen that the eight-parameter model had the lowest AIC score. This model has a fixed knee point position and symmetrical horizontal declines, but asymmetrical superior and inferior declines. This choice was also supported by nested-model hypothesis testing (F-statistic, with significance set at $p = 0.05$). The RMS error of this fit is only marginally (0.02 dB) worse than that provided by fitting each meridian separately (top row of Table 4.3).

	Superior		Inferior		Horizontal		Global	
Observer	m_1	m_2	m_1	m_2	m_1	m_2	ν	k_2
ASB	1.05	0.40	1.05	0.21	0.76	0.31	8.0	8.6
DHB	1.05	0.56	1.31	0.27	1.00	0.47	8.2	12.6
SAW	1.01	0.49	1.09	0.33	0.92	0.44	10.5	12.2
TSM	1.38	0.49	1.35	0.28	1.22	0.45	7.9	8.5
Average	1.12	0.50	1.20	0.28	0.97	0.43	8.5	10.5

Table 4.4: Parameters obtained from fitting the preferred (eight-parameter) witch's hat bilinear model (see Table 4.3) to the combined data from Experiments 1 and 2. Different rows show the parameters given by fitting to each observer's data individually, and to the average of the observers' data.

The eight-parameter model fits to the combined data from Experiments 1 and 2 (averaged across all observers) are shown in Figure 4.5. Table 4.4 shows the parameters for these fits, and for fitting the same model to the individual observer data. The global ν parameter for this model means that the knee point location is constant across different meridians (at 8.5 cycles for the averaged observer). Fitting where ν was allowed to vary with the m_1 and m_2 parameters (i.e. different knee point locations were permitted for superior, inferior and horizontal) neither improved the quality of the fit nor changed the value of the fitted ν parameter.

4.7.4 Radial interpolation

The bilinear model has been shown to provide a good fit to the experimental data collected along the cardinal meridians, but the desired outcome for this series of experiments is to generate a map of contrast sensitivity across the central visual field. Abrams et al. (2012) found the sensitivity along diagonal ("intercardinal") meridians to be intermediate between the sensitivities at the cardinal meridians. This suggests that predictions for sensitivity along diagonal meridians can be made via interpolation between the fitted cardinal meridians to create a 2D attenuation surface. This can be achieved by varying the fitted bilinear model parameters elliptically with polar angle (see MATLAB code in Section B.2).

The radial interpolation method was tested using the results from Experiment 1. An eight-parameter bilinear fit was made to the data from the cardinal meridians (horizontal and vertical). These fits were then interpolated to produce predictions for the two diagonal meridians. A comparison of this interpolated prediction against a direct fit to the data (shown in Figure 4.8, see also the caption to this figure) provides evidence that this method is valid.

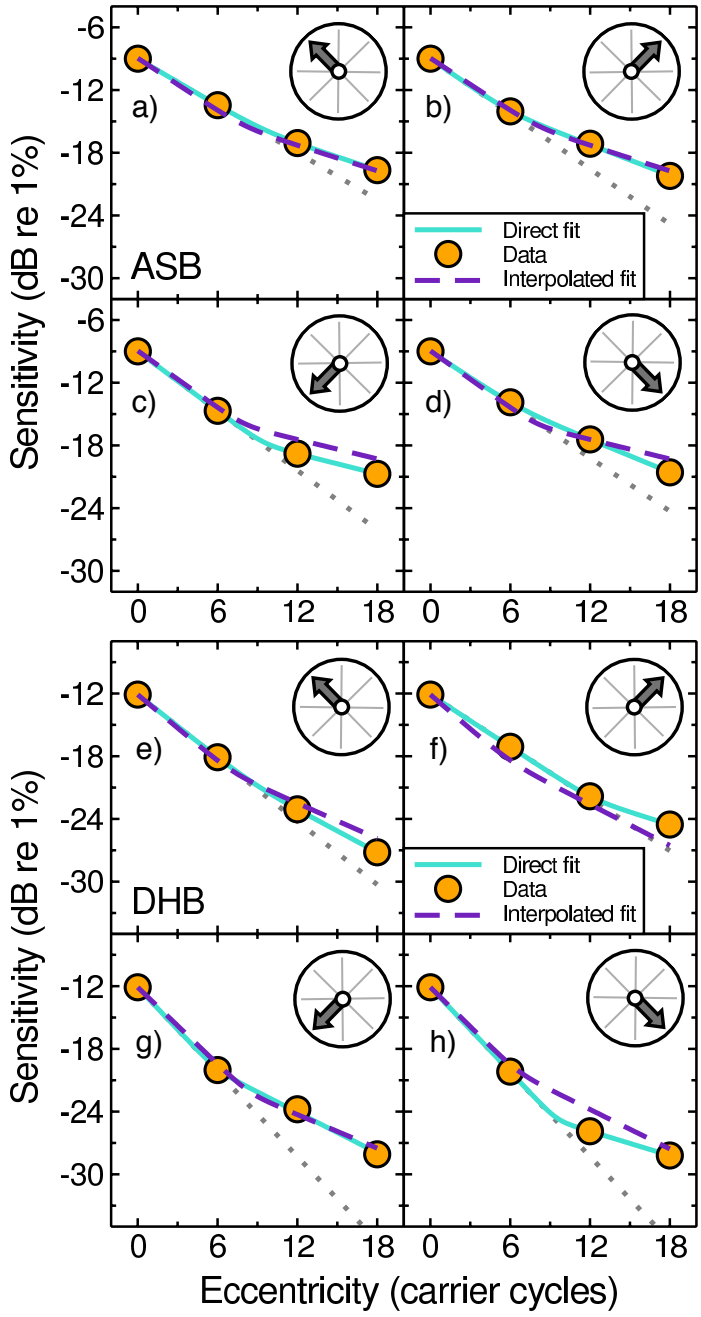

Figure 4.8: A comparison of a direct fit to the data collected along diagonal meridians in Experiment 1 against a prediction made by interpolating between the fitted horizontal and vertical meridians. The left and right plots show the data for ASB and DHB respectively. The direct fit is provided by fitting each diagonal meridian individually with a four-parameter witch's hat bilinear function, allowing the gradients (m_1 and m_2) and knee point position (ν) to vary but fixing the vertical offset (k_2) across all four hemi-meridians (giving 13 free parameters per plot). The interpolated fit was created by generating a surface using just the horizontal and vertical data, to then extract the radially interpolated diagonal values from that surface (no free parameters). For ASB the RMS errors for the direct and interpolated fits were 0.03 and 0.63 dB respectively. For DHB the RMS errors were 0.09 and 0.96 dB. The dotted grey lines are extrapolations of the gradient of the initial decline from the direct fit (m_1).

4.7.5 Scale invariance

Previous studies have found the decline in contrast sensitivity to be scale invariant within broad spatial frequency ranges. Pointer and Hess (1989) found that, for spatial frequencies of 1.6 to 12.8 c/deg, the log-sensitivity versus eccentricity functions were vertical translations of each other when eccentricity was expressed in stimulus carrier cycles. Although this may be true for the linear declines in log sensitivity reported previously, it is possible that the more complex bilinear decline reported here violates this scale invariance. For example it could be that the knee point location (ν), rather than being located at an eccentricity which is constant when expressed in carrier cycles (8.5 cycles for the fit to the averaged data), is located at a constant position on the retina (2.125 degrees).

Parameter	RMS error for model variant (dB)							
ν	Fix ν in cycles				Fix ν in degrees			
m_1	m_1 dB/cycle		m_1 dB/degree		m_1 dB/cycle		m_1 dB/degree	
m_2	m_2 cy	m_2 deg	m_2 cy	m_2 deg	m_2 cy	m_2 deg	m_2 cy	m_2 deg
ASB	0.740	0.827	1.143	2.258	1.099	1.726	1.252	2.359
DHB	1.157	1.185	0.977	1.877	0.757	2.142	1.007	1.938
SAW	0.679	1.322	1.744	3.000	1.233	9.380	1.445	3.272
Average	0.594	1.238	0.864	2.290	0.745	1.784	0.981	2.422

Table 4.5: RMS errors for the 2x2x2 factorial model analysis. This covers each possible combination of the two methods of scaling the model parameters (i.e. whether they are fixed across spatial frequency in terms of carrier cycles or in terms of visual angle). Each model variant was fit to individual datasets from Experiment 3 for each observer (the models in the leftmost and rightmost columns of this table are shown in the left and right columns of Figure 4.6 respectively), and to data averaged across the three observers. The models were fit to the data from all spatial frequencies simultaneously in each case. For space reasons, the "dB/cycle" and "dB/degree" units in the row of the table defining the behaviour of the m_2 parameter are abbreviated to "cy" and "deg" respectively.

The data from Experiment 3 were fitted with several different variations of the bilinear model. This tested each combination of a parameter being fixed in terms of degrees, or in terms of cycles. Doing this for both ν (testing whether the location of the knee point was a constant number of cycles across spatial frequency or whether its location was a fixed number of degrees from fixation) and for the slope parameters m_1 and m_2 (testing whether the gradients of the decline were a constant number of decibels per cycle or per degree of visual angle) resulted in a comparison between eight different models which are listed in Table 4.5.

The data for all three observers (and their averaged data) are better fit by a model where all parameters are defined in cycles rather than in degrees (compare the column on the left of Table 4.5 against the rightmost column). For ASB, SAW and the averaged data this fit was also superior to all of the other combinations. DHB showed a preference to a fit where the knee point was defined in terms of degrees of visual angle (rather than carrier cycles), with the gra-

dients still being defined in dB/cycle. The RMS error of the scale invariant fit for DHB is not much worse than for this preferred model however, so it is concluded that a scale-invariant description of the contrast sensitivity decline is adequate within the tested range.

Observer	Gradients (dB/cycle)				Knee (cycles)	
	m_1		m_2		ν	
ASB	1.14		0.40		7.94	
DHB	0.97		0.55		9.85	
SAW	2.15		0.63		4.64	
Average	1.08		0.40		10.5	
	Vertical offset k_2 (dB) per spatial frequency					
	0.7 c/deg	1 c/deg	1.4 c/deg	2 c/deg	2.8 c/deg	4 c/deg
ASB	-5.40	-3.50	-5.49	-4.89	-8.45	-9.66
DHB	-5.94	-6.30	-6.79	-9.18	-11.11	-14.16
SAW	-0.73	-1.27	-1.91	-6.21	-5.29	-9.90
Average	-4.61	-4.28	-5.32	-7.35	-8.87	-11.82

Table 4.6: Parameters for the witch's hat bilinear model fits to the data from Experiment 3 for each observer (plotted in the left column of Figure 4.6) and to the data averaged across the three observers.

The parameters for the scale invariant fits (i.e. those from the leftmost column of Table 4.5) are provided in Table 4.6. These are shown fitted to the data in the left column of Figure 4.6 (the right column shows fits where parameters were defined in terms of degrees of visual angle). The parameters for optimal fits for the three observers tested here show greater variability than was present in the fits to the data from Experiments 1 and 2 (Table 4.4). This is partially due to the interaction between the gradient parameters and the location of the knee point. Fixing the position of the knee point for each observer such that it is constrained to be in the same location as was reported in Table 4.4 provided similar scale invariant fits for a small increase in RMS error (not shown).

4.7.6 Comparison with physiology

A scale-invariant decline in contrast sensitivity is in agreement with previous psychophysical results (Robson & Graham, 1981; Pointer & Hess, 1989). Some previous studies have however made use of attenuation surfaces based on retinal physiology (e.g. Ahumada & Watson, 2011). A physiologically-derived model was created in order to investigate whether scale invariance could be predicted from the decline in photoreceptor density with eccentricity. Cone density data were extracted from Curcio et al. (1990) and fitted with a polynomial function. This description of density as a continuous one-dimensional function of eccentricity (Figure 4.9), was then radially interpolated to create a two dimensional function (d). This density map was used

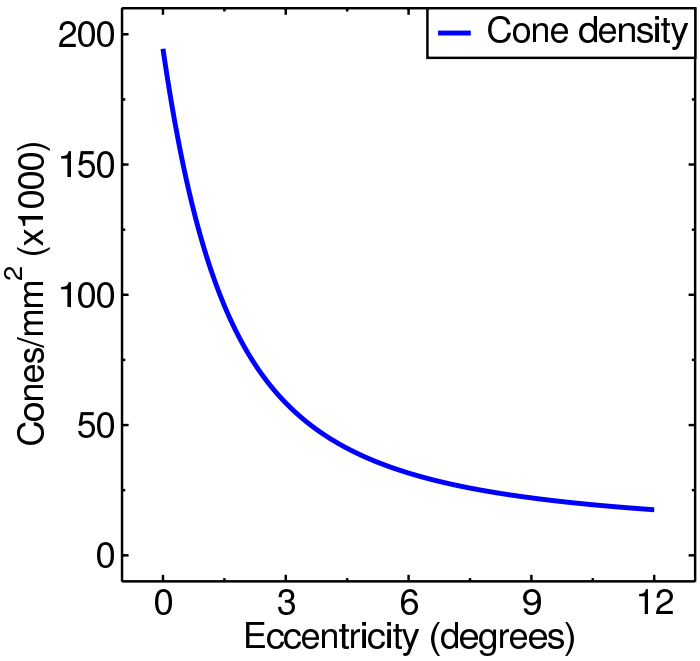

Figure 4.9: A description of the decline in cone density (in thousands of cones/mm^2) with eccentricity, derived by fitting a 10th order polynomial to the density data of Curcio et al. (1990). Data were combined over the superior, inferior, nasal and temporal hemi-meridians. This one-dimensional function is radially interpolated to make the two-dimensional function d, used in the modelling here.

to simulate a square cone matrix, after Anderson et al. (1991). A contrast attenuation factor (a), which is proportional to the Nyquist frequency (Anderson et al., 1991; Ahumada & Watson, 2011), was calculated from the cone density as

$$a_i = \sqrt{d_i}. \tag{4.7}$$

This function was then normalised to have a gain of unity at fixation.

The attenuation surface derived from the cone density function was incorporated into a model that combined responses from individual locations over space in order to predict the relative sensitivity (presented as the signal-to-noise ratio or SNR) to log-Gabor patches at different eccentricities (0 - 12 degrees) and spatial frequencies (0.7 - 8.0 c/deg).

The sensitivity predictions were calculated as follows. The response (r_i) at each (i) of n locations is weighted by the attenuation factor for that location (a_i). Matching a template to the expected signal results in these terms being squared in the numerator, after which they are scaled by the stimulus contrast (c). I assume the signal at each location to be perturbed by independent Gaussian noise with unit standard deviation. Therefore, the noise is also weighted by the template. The standard deviation of the noise at each location is squared, and then these variances are summed and the square root is taken to find the standard deviation of the

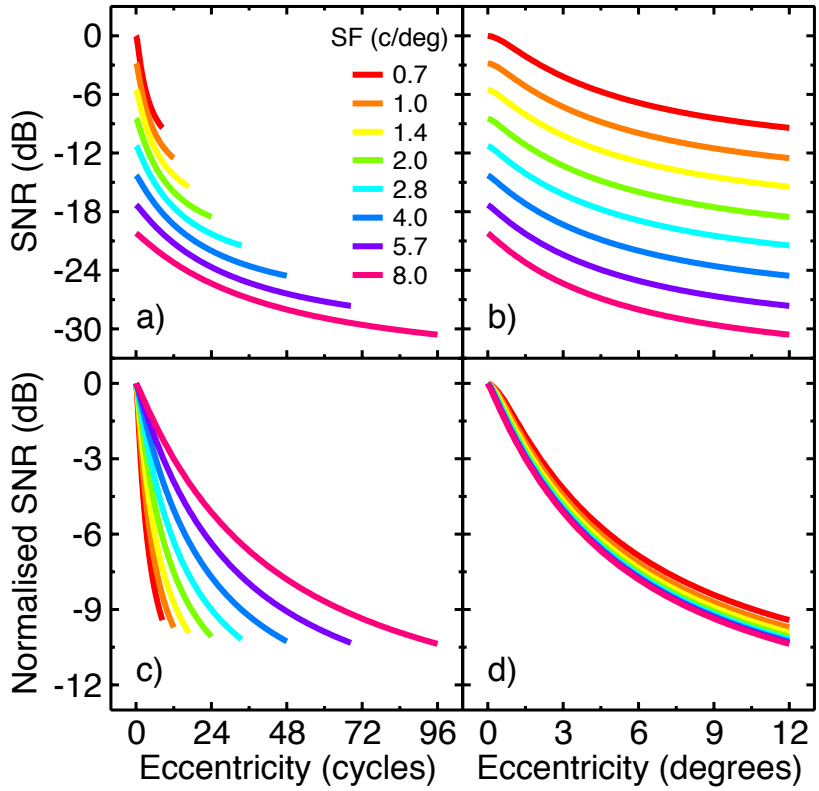

Figure 4.10: Relative contrast sensitivity predicted by the signal-to-noise ratios derived from Equation 4.9 and the cone density data in Figure 4.9. Eccentricity is plotted in carrier cycles in the left column, and in degrees of visual angle in the right column. In the top row each curve is normalised to the SNR for the 0.7 c/deg patch at fixation. In the bottom row each curve is normalised to SNR for that curve at fixation.

combined noise. The signal-to-noise ratio (SNR) is therefore calculated as

$$\text{SNR} = \frac{\sum_{i=1}^{n}(cs_i^2a_i^2)}{\sqrt{\sum_{i=1}^{n}(s_i^2a_i^2)}}. \tag{4.8}$$

Simplifying Equation 4.8 for a contrast of unity and substituting $\sqrt{d}$ for a (Equation 4.7) gives the equation that I used to create the curves shown in Figure 4.10

$$\text{SNR} = \sqrt{\sum_{i=1}^{n}(s_i^2d_i)}. \tag{4.9}$$

The curves in Figure 4.10 are close to parallel when eccentricity is expressed in degrees, but diverge when it is expressed in carrier cycles. The sensitivity functions expressed against eccentricity in carrier cycles are simply those for the case where eccentricity is expressed in degrees scaled horizontally by a factor equal to the spatial frequency, these two properties are

fundamentally incompatible. The separation between the normalised curves in Figure 4.10d is due to the rounding off of the peaked sensitivity decline function caused by the larger absolute size of the stimuli for the lower spatial frequencies (the effect of using stimuli of any size in this model is effectively to convolve the sensitivity decline with the stimulus envelope).

The prediction of the attenuation surface based on cone density shown here is most similar to the scale invariant result when the spatial frequency of the stimulus is approximately 2 c/deg. This may account for why previous models that have used an attenuation surface based on cone density could provide acceptable fits for data from stimuli in that spatial frequency range (e.g. Ahumada & Watson, 2011). The lack of scale invariance demonstrated with the cone density model shown here does not depend critically on either the precise shape of the decline or how the signal and noise are combined over space. Altering these factors produced similar behaviour, merely changing the slope or the shape of the scale-dependent decline.

4.8 Further experiments in the periphery

4.8.1 The sensitivity decline from 18 to 62 cycles

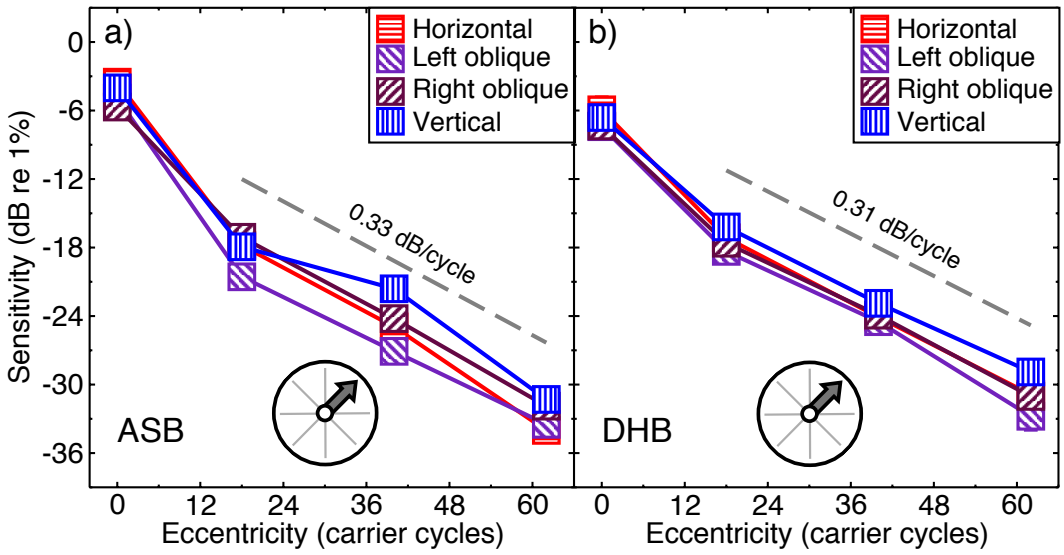

Figure 4.11: Data from two observers (ASB and DHB) showing the decline in contrast sensitivity along the 45° (upper-right diagonal) hemi-meridian to the four different patch orientations. The target spatial frequency was 4 c/deg. These data were collected at greater eccentricities than those tested in Experiment 1. The solid grey lines show linear fits to the most eccentric three data points (those at: 18, 40 and 62 cycles; equivalent to 4.5, 10, and 15.5 degrees).

To reconcile this study with the results of Sasaki et al. (2006) who *did* find a strong relative orientation effect, an additional experiment was carried out where more eccentric locations

along the 45° (superior-right diagonal) hemi-meridian were tested (0, 18, 40 and 62 cycles; equivalent to 0, 4.5, 10 and 15.5 degrees) using larger stimuli to bring the contrast thresholds for these locations down into a measurable range (the spatial frequency and orientation bandwidths were halved to 0.8 octaves and ±12.5° respectively). There is no evident relative orientation effect present in the results (Figure 4.11), which for this hemi-meridian would cause a higher sensitivity to the right-oblique patches. This is the case even at 15.5 degrees of visual angle (the eccentricity tested by Sasaki et al., 2006). Sensitivity appears to decrease with eccentricity in the same manner for all four patch orientations, undergoing a slightly steeper decline within the first 18 cycles (due to the m_1 region) and then transitioning to a shallower slope that continues out to at least 62 cycles (0.33 and 0.31 dB/cycle respectively for the observers ASB and DHB here). These gradients are slightly shallower than both the directly fitted m_2 values (0.49 and 0.39 dB/cycle) and the interpolated m_2 values (0.41 and 0.61 dB/cycle) for this hemi-meridian (shown in Figure 4.8).

4.8.2 Relative orientation effects appear on the horizontal meridian

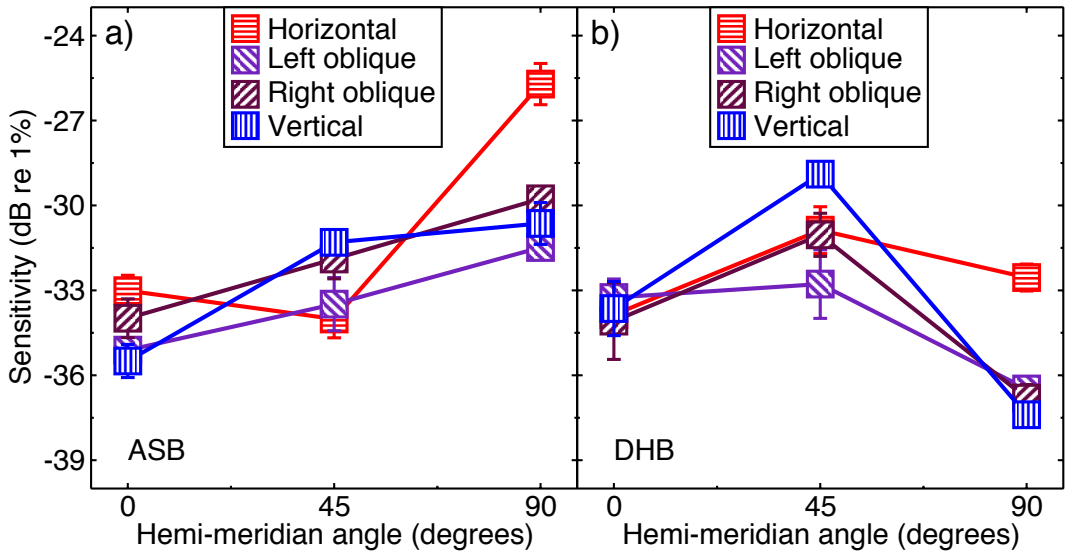

Figure 4.12: Data from two observers (ASB and DHB) showing the contrast sensitivity to the four different patch orientations at an eccentricity of 62 cycles (15.5 degrees), as a function of hemi-meridian angle. The data for the 45° hemi-meridian are replotted from Figure 4.11.

Additional data were collected at the greatest eccentricity (62 cycles, equivalent to 15.5 degrees) for the 0° and 90° hemi-meridians (superior-vertical and right-horizontal). The contrast sensitivity for each patch orientation is plotted as a function of hemi-meridian angle in Figure 4.12. A relative orientation effect is visible for both observers on the 90° hemi-meridian, as evidenced by large advantage in the detectability of horizontal patches at this location (4.9 and 4.3 dB for observers ASB and DHB respectively). The two other hemi-meridians do not

show this effect. Testing the 90° hemi-meridian at a lesser eccentricity of 40 cycles (10 degrees) did not find any advantage of horizontal stimuli over vertical stimuli, indicating that the relative orientation effect does not appear before this point (i.e. it is restricted to the more eccentric locations).

It is possible that relative orientation effects for other meridians might also appear at greater eccentricities than those tested here, however for the purpose of this study (seeking to characterise the central visual field) the results indicate that these effects can be ignored. It is also not possible to conclude from these results whether the effect is caused by neural factors or if it is due to optical aberrations in the eye. An uncorrected cylindrical refractive error in the periphery would selectively affect the detection of stimuli with particular orientations, dependent on the axis of the cylinder. This confound was ruled out as an explanation of the resolution effects found by Rovamo et al. (1982), by attempting to correct any possible refractive error with cylinder lenses. The previous study that reported a contrast sensitivity effect however did not control for these optical factors (Sasaki et al., 2006).

4.9 Discussion

4.9.1 Bilinearity

Rather than the linear decline in log contrast sensitivity with eccentricity described in the previous literature, it is demonstrated here that the shape of the decline is bilinear within the central visual field (for stimuli within the spatial frequency range 0.7 - 4 c/deg). The initial slope of this decline (0.97 - 1.20 dB/cycle) is more than twice as steep as slopes reported previously for linear declines (0.33 - 0.50 dB/cycle in Pointer & Hess, 1989). This indicates that stimuli shown in the centre of the visual field are subject to an attenuation which is more severe than would be suggested from previous results. The bilinear decline transitions to a shallower slope at a knee-point which is placed at approximately 8.5 cycles, after which it has a gradient of less than half its original value (0.28 - 0.50 dB/cycle).

The rate of decline in contrast sensitivity after the knee point is much closer to the classical results (e.g. Pointer & Hess, 1989), which is likely due to those studies measuring over a wider area and therefore biasing their results to the secondary slope. There is some evidence of bilinearity in the data of Pointer and Hess (1989), however their relatively sparse sampling of eccentricities acts to obscure this feature. It is also possible that the larger stimulus sizes used in previous studies would act to blur the features of the decline. In agreement with previous results however, the data collected from a wider range of eccentricities (shown in Figure 4.11) suggest that after the knee point the decline continues at roughly the m_2 gradient out to at least 62 cycles.

4.9.2 Scale invariance

The scale-invariant decline reported here is consistent with previous results (e.g. Robson & Graham, 1981; Pointer & Hess, 1989), though a scale invariant decline in contrast sensitivity can not be easily derived from retinal physiology (Section 4.7.6). It is unclear how the scale invariance I find here can be accounted for (though see Section 4.9.5), however the attenuation surface described here is accurate to the psychophysical data which is the goal of this study.

4.9.3 Orientation effects

The experiments that investigated the oblique effect found only small and inconsistent results. The relative orientation effect found at 62 cycles confirms the results of Sasaki et al. (2006), though the experiments here indicate that strong effects are confined to the more eccentric locations ($>$40 cycles). The absence of orientation effects (relative or absolute) within the central visual field indicates that these are not relevant to the purposes of this study. This means that a map of the sensitivity in this area can be derived using stimuli of a single orientation.

4.9.4 Meridional anisotropies and the attenuation surface

The findings from this study confirm the two meridional anisotropies reported in the previous literature (the "horizontal-vertical" and "superior-inferior" anisotropies reviewed by Abrams et al., 2012). The decline in sensitivity along the horizontal meridian is shallower than that along the vertical meridian (in agreement with Pointer & Hess, 1989). The relative steepness of the initial (m_1) decline along the superior and inferior hemi-meridians varied across the observers. For ASB and SAW the slopes of the decline were roughly equal for the superior and inferior hemi-meridians; whereas for DHB, TSM and the averaged data the initial decline along the inferior hemi-meridian was steeper than that along the superior hemi-meridian. For all four observers however (and for their averaged data) the second (m_2) limb of the bilinear function was shallower along the inferior than along the superior hemi-meridian. This is in agreement with the previously reported "superior-inferior" anisotropy (Abrams et al., 2012).

The success of interpolating the fits to the cardinal meridians to predict the sensitivity along the diagonal meridians indicates that it is only necessary to collect data along the horizontal and vertical meridians to generate an attenuation surface. The parameters being mirrored across the horizontal meridian means that the amount of data collected along this meridian can be halved whilst still providing a prediction with the same level of confidence for this meridian as for the vertical meridian (by collapsing data across the horizontal hemi-meridians). The scale-invariance of the result within the frequency range tested (0.7 - 4 c/deg) allows an attenuation surface measured for one spatial frequency to be generalised to other spatial frequencies within this range (when eccentricities are expressed in stimulus carrier cycles).

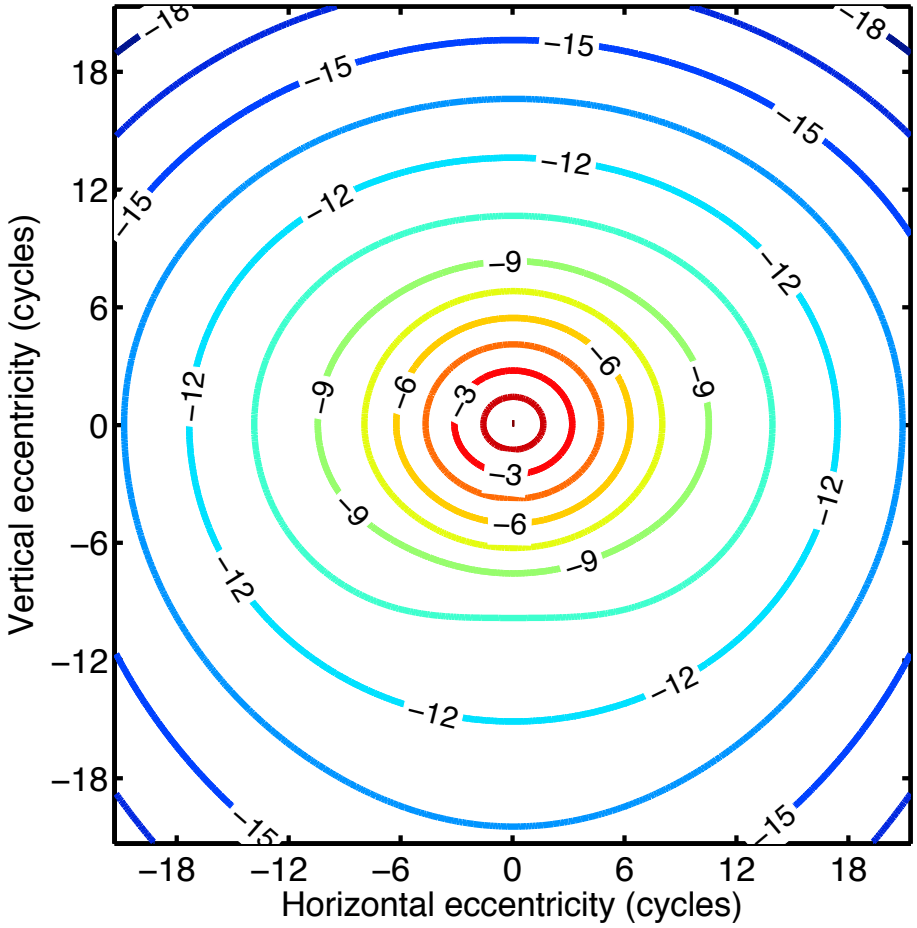

Figure 4.13: A contour map of contrast sensitivity within the central 18 cycles of the visual field. The numbers labelling the contours indicate the amount of contrast attenuation (in dB). This map is based on a radial interpolation of the model fit to the average data in the bottom row of Table 4.4.

The witch's hat generated by radially interpolating the fits to the averaged data from Experiment 2 (the bottom row of Table 4.4) is shown as a contour plot in Figure 4.13 and as a three-dimensional surface (two spatial dimensions and contrast sensitivity) in Figure 4.14. The MATLAB code for producing the attenuation surface is provided in Section B.2 of Appendix B. This attenuation surface gives a more detailed account of the inhomogeneity in contrast sensitivity within the central visual field than was previously available. It is useful both as a stage in a model of spatial vision (see Section 5.2.1) and to transform stimuli before presentation in order to make them equally detectable at each point in space (see Chapters 6 and 7).

4.9.5 Witch hat compensation

Other studies presented in this thesis (Chapters 6 and 7) provide a proof of concept that applying witch hat compensation for the visual field inhomogeneity can factor out the effects of eccentricity on effective stimulus contrast for area summation tasks. Previous attempts have

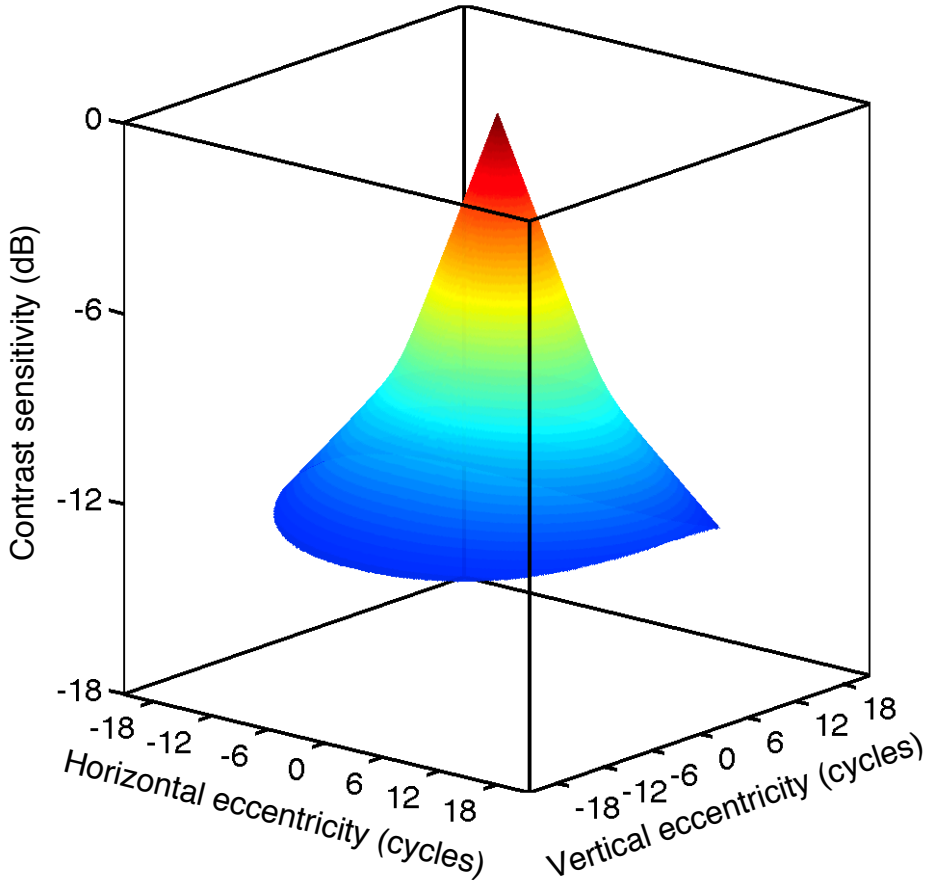

Figure 4.14: The same map of the contrast sensitivity inhomogeneity as that shown in Figure 4.13, displayed as a three dimensional surface. It is from the shape of this surface that the "witch's hat" bilinear model draws its name.

been made to equalise performance across the visual field by increasing the size of a stimulus with eccentricity, either based on the cortical magnification factor (Daniel & Whitteridge, 1961; Rovamo et al., 1978; Rovamo & Virsu, 1979; Virsu & Rovamo, 1979) or on empirical psychophysical results (Watson, 1987). Rovamo and Virsu (1979) applied this M-scaling method and successfully brought contrast sensitivity functions measured in different regions of the visual field into a reasonable correspondence, however they fell short of equalising contrast thresholds (by a factor of 2). Several other studies have also shown that size scaling alone does not equalise performance (for reviews see García-Pérez & Sierra-Vázquez, 1996; Strasburger et al., 2011). More recent work has overcome these failings by building contrast scaling into the compensation alongside size scaling (Mäkelä, Näsänen, Rovamo, & Melmoth, 2001; Rovamo & Melmoth, 2002; Melmoth & Rovamo, 2003). For the contrast detection threshold experiments conducted in this thesis, contrast scaling alone appears to be sufficient. It is possible however that this could be replaced by a combination of contrast and size scaling.

Although the attenuation surface provides an accurate description of how the empirical contrast sensitivity varies with eccentricity, the functional properties of the visual system that it

models are open to interpretation. The simplest architecture implied by the summation models used in this thesis (see Chapter 5) is that of a visual field tiled by a regular square array of detectors. In this case the function of the witch's hat surface as applied in the model is to increase the gain of the less eccentric detectors relative to that of the eccentric detectors (an alternative account would be that the internal noise becomes greater with eccentricity). Based on what is known about the cortical magnification factor in the mapping of the visual field by the brain however, the assumption of a regular array of detectors with equal density spanning the visual field seems unlikely.

A more likely account of the witch's hat attenuation surface consistent with the idea of an inhomogeneous mapping of the visual field would be if detectors tuned to each spatial frequency decreased in density with eccentricity (with a common rule based on eccentricity in cycles, see Section 4.7.5). If detectors were affected by independent Gaussian noise, each location would have sensitivity proportional to the square root of the number of detectors at that location. The decline in performance with eccentricity could then be accounted for by a concomitant reduction in the density of detectors tuned to the stimulus. Under this account, the attenuation surface as applied in the modelling here would simply be a shorthand for a more complicated process of performing spatial transformations on the stimulus to decrease the sampling density in the periphery.

CHAPTER 5

Summation modelling

5.1 Introduction

5.1.1 Analytic and stochastic models of area summation

This chapter presents general models developed to account for the summation of contrast over area. There are three types of model described here: analytic, stochastic and analytic approximation. Analytic models feature a closed-form solution for the detection threshold of a stimulus based on the signal-to-noise ratio (SNR; see Section 3.2.6) at the decision stage. Stochastic models simulate a visual system with the modelled architecture and behaviour responding to the experimental task, and analyse the data from these simulations to generate model predictions. Analytic approximations feature closed-form shorthands that approximate the behaviour of models for which the analytic solution would be difficult or impossible to derive. In the application of these models in future chapters only analytic versions and analytic approximations are used where possible, however stochastic versions were developed alongside these to ensure their behaviour was identical and to model conditions for which no analytic solution or approximation was available.

5.1.2 Fitting summation models to data

Where predictions from summation models are presented in this thesis they are fitted to the data to minimise the RMS error of the fit using a Simplex algorithm (see Section 3.7). The fits are performed by allowing the model prediction curves to translate vertically along the log-threshold axis. For all models, this corresponds to a global change in the sensitivity of the modelled visual system (i.e. the ratio between the magnitude of the input signal and the standard deviation of the limiting noise). In this sense the parameter that varies in order to fit the models to the data does not reflect any changes in the summation behaviour of those models.

5.2 Model stages

5.2.1 Stimulus attenuated according to contrast sensitivity inhomogeneity

Figure 5.1: Multiplication of the "Battenberg" stimulus in panel a) (see Section 3.4.6) by the witch's hat attenuation surface in panel b) (see Section 4.9.4) gives the attenuated stimulus in panel c). The contrast of the image decreases from the centre outward at the same rate as the decline in contrast sensitivity in human vision.

In all of the models featured here, the contrast of the stimulus image is first adjusted to reflect the variation in sensitivity across the visual field (see Figure 5.1). A witch's hat attenuation surface derived from empirical measurements (see Section 4.9.4) is generated for the appropriate observer, fixation location, and spatial frequency. This is a 2D matrix the same size as the stimulus image with the value at each location being the gain of the input stage of the visual system at that location relative to that at fixation (giving a gain of unity at fixation, and below that elsewhere). This attenuation surface ($\mathbf{A}$) is multiplied by the stimulus image ($\mathbf{S}$) to provide the attenuated image ($\mathbf{S}_{\text{att}}$), which is used as the input for the next stage of the model

$$\mathbf{S}_{\text{att}}[x, y] = \mathbf{S}[x, y] \cdot \mathbf{A}[x, y]. \tag{5.1}$$

5.2.2 Spatial filtering by log-Gabor patches

The next stage is the spatial filtering of the image in analogy to the process performed by the visual system up to the simple cell stage in V1 (see Section 2.3.1). Where stimuli were presented with a single target orientation and spatial frequency (as is the case for all of the contrast detection studies presented here) this is modelled as occurring within a single orientation- and frequency-tuned channel (see Section 2.3.2). This is implemented by convolving the stimulus images with a log-Gabor patch.

The bandwidths of the log-Gabors used here ($\pm 25^\circ$ orientation, 1.6 octaves spatial frequency) were consistent with those used in previous summation models, chosen there to reflect the

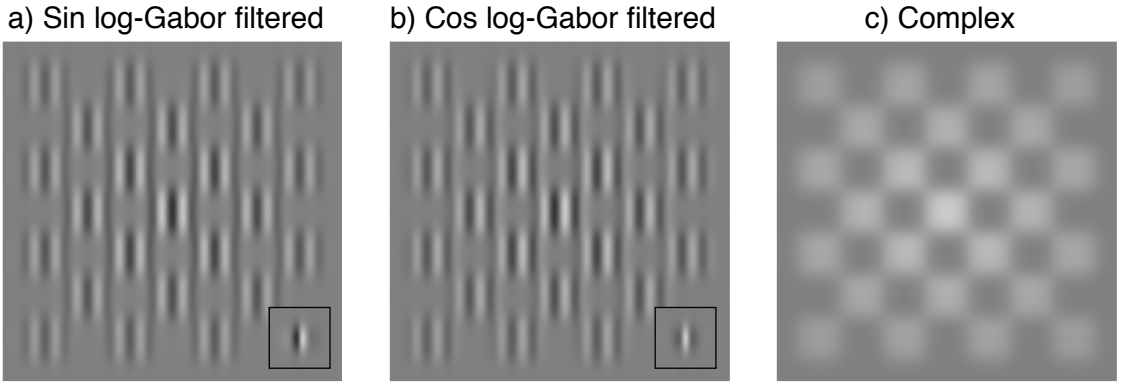

Figure 5.2: The output of filtering the attenuated stimulus (Figure 5.1c) with a sine-phase log-Gabor (a) and a cosine-phase log-Gabor (b). The simulated complex cell response calculated by taking the Pythagorean sum of the sine and cosine responses is shown in c). Inset in a) and b) are the sine-phase ($\mathbf{L}_{\text{sin}}$) and cosine-phase ($\mathbf{L}_{\text{cos}}$) log-Gabors used to perform the filtering.

bandwidths found in simple cells (Meese & Summers, 2007; Meese, 2010). The output of this stage is an image where the intensity of each pixel reflects the activity of a model simple cell (also referred to as a "detector") at that location (see Figure 5.2). The spatial frequency, orientation, and phase tuning is defined by the the properties of the log-Gabor filter element.

The responses of the sin-phase log-Gabor filter elements ($\mathbf{S}_{\text{sin}}$) are calculated by convolving the attenuated stimulus image ($\mathbf{S}_{\text{att}}$) with a sin-phase log-Gabor ($\mathbf{L}_{\text{sin}}$)

$$\mathbf{S}_{\text{sin}} = \mathbf{S}_{\text{att}} * \mathbf{L}_{\text{sin}}, \tag{5.2}$$

the cosine-phase responses ($\mathbf{S}_{\text{cos}}$) are calculated using a cosine-phase log-Gabor ($\mathbf{L}_{\text{cos}}$)

$$\mathbf{S}_{\text{cos}} = \mathbf{S}_{\text{att}} * \mathbf{L}_{\text{cos}}, \tag{5.3}$$

and the complex response ($\mathbf{S}_{\text{complex}}$) is calculated from the Pythagorean sum of the sine and cosine responses

$$\mathbf{S}_{\text{complex}}[x, y] = \sqrt{\mathbf{S}_{\text{sin}}[x, y]^2 + \mathbf{S}_{\text{cos}}[x, y]^2}. \tag{5.4}$$

For the summation modelling presented here there is little to no difference in predictions made by models with sine phase, cosine phase, or complex responses from the filtering stage.

To perform the convolution, the Fourier transform of the attenuated image is multiplied by the Fourier transform of a log-Gabor patch. The output of this process is then converted back to the spatial domain to give the filtered image. Summation within the simulated receptive fields of this stage bypasses any subsequent nonlinearities. This within-filter summation causes models to behave as if they were linear for stimuli which are smaller than the filter elements, and increases the predicted summation for stimuli of a similar size to the filter element.

5.2.3 Rectification and nonlinear transduction of filter outputs

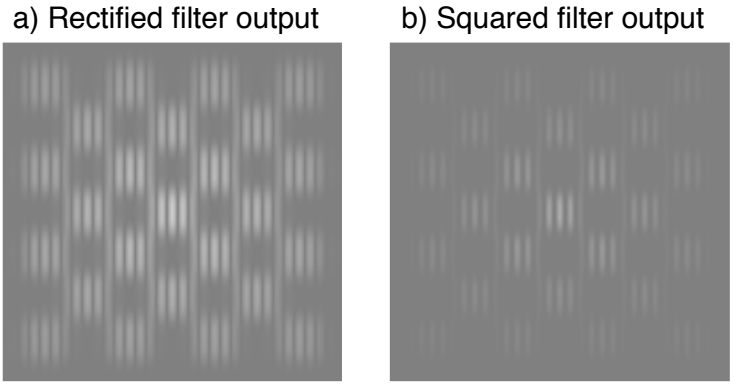

Figure 5.3: The rectified output from filtering with a sin-phase log-Gabor patch (Figure 5.2) is shown in panel a). The image in panel b) shows the effect of squaring the value at each pixel (representing the nonlinear transduction of filter outputs).

The filter outputs are rectified ($\mathbf{S}_{\text{rect}}$) by taking the absolute value of each pixel (see Figure 5.3a)

$$\mathbf{S}_{\text{rect}}[x, y] = |\mathbf{S}_{\text{sin}}[x, y]|. \tag{5.5}$$

This represents the unsigned magnitude of the filter outputs.

The pixel values representing the filter outputs may then undergo nonlinear transduction ($\mathbf{S}_{\text{trans}}$) by raising them to a power m (see Figure 5.3b)

$$\mathbf{S}_{\text{trans}}[x, y] = (\mathbf{S}_{\text{rect}}[x, y])^m. \tag{5.6}$$

5.2.4 Pixelwise additive Gaussian noise

The output of each filter element is perturbed by internal noise. This is modelled as independent additive Gaussian noise with a mean of zero and a constant variance. The assumption of Gaussian noise is made in accordance with the Central Limit Theorem (Peterson et al., 1954; Tyler & Chen, 2000). In stochastic models (see Figure 5.4a), this is simulated and added to the transduced stimulus image to give the noisy filter outputs ($\mathbf{S}_{\text{noisy}}$)

$$\mathbf{S}_{\text{noisy}}[x, y] = \mathbf{S}_{\text{trans}}[x, y] + N(\mu, \sigma^2), \tag{5.7}$$

where $N(\mu, \sigma^2)$ is a sample from Gaussian noise with mean μ and standard deviation σ. For analytic models (see Figure 5.4b), the noise is represented in the calculations by a separate matrix ($\mathbf{G}$) containing the standard deviations of the noise for the output of each filter element

$$\mathbf{G}[x, y] = \sigma. \tag{5.8}$$

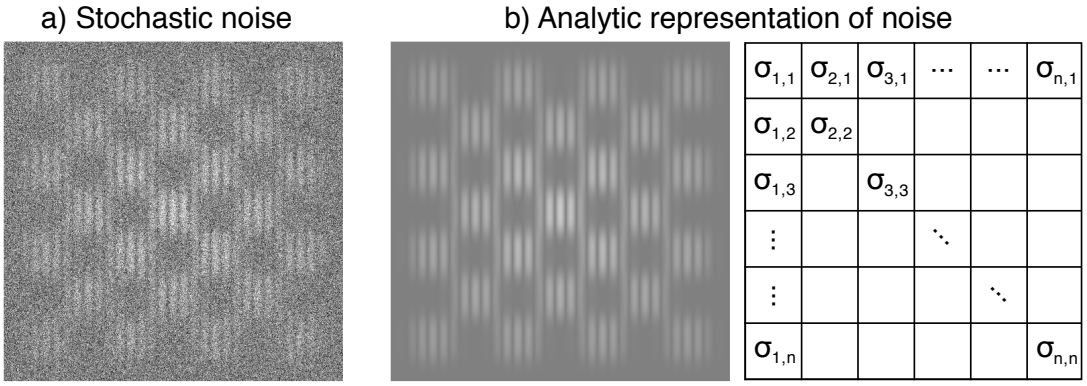

Figure 5.4: Noise is represented differently in the two types of model. In stochastic models (a) independent Gaussian noise is added to the pixel value at each location in the filtered stimulus image. In analytic models (b) the noise is represented as a separate matrix containing the standard deviations of the noise for each pixel in the filtered stimulus.

The only model architectures considered here are those where the dominant source of noise comes after the nonlinear transduction stage. According to Birdsall's theorem, dominant noise placed *before* transduction linearises the transducer. This makes the behaviour of such a system equivalent to that of a system with a linear transducer (though see Appendix A).

5.2.5 Template matching

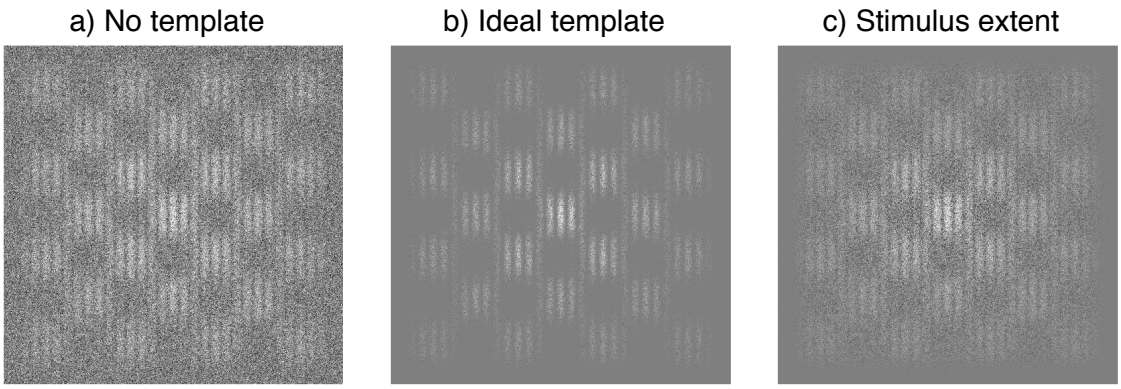

Figure 5.5: Different template strategies are shown here, demonstrated using the stochastic model. Panel a) shows the noisy stimulus image with no template applied. Panel b) shows the image multiplied by a template which is matched to the stimulus exactly (an "ideal" template). Panel c) shows the image multiplied by a template that is matched to the stimulus extent (without the "Battenberg" modulation). In both b) and c) the weighting of the templates declines with eccentricity in proportion to the expected signal to noise ratio resulting from the attenuation surface.

An observer behaving with knowledge of the expected stimulus could choose to improve the signal-to-noise ratio at the decision stage by combining a weighted input from each detector, according to a template. The "ideal" template would be matched exactly to the stimulus (see

Figure 5.5b). The output of this stage ($\mathbf{S}_{\text{temp}}$) is obtained by multiplying the noisy stimulus by the attenuated, filtered, and transduced stimulus ($\mathbf{S}_{\text{trans}}$) in the stochastic model

$$\mathbf{S}_{\text{temp}}[x, y] = \mathbf{S}_{\text{noisy}}[x, y] \cdot \mathbf{S}_{\text{trans}}[x, y], \tag{5.9}$$

whilst in the analytic model the template is applied both to the signal matrix and to the standard deviations of the pixelwise noise

$$\mathbf{S}_{\text{temp}}[x, y] = \mathbf{S}_{\text{trans}}[x, y] \cdot \mathbf{S}_{\text{trans}}[x, y], \tag{5.10}$$

$$\mathbf{G}_{\text{temp}}[x, y] = \mathbf{G}[x, y] \cdot \mathbf{S}_{\text{trans}}[x, y]. \tag{5.11}$$

Models with a template matched to the stimulus envelope make similar predictions to those made by a model with the ideal matched template. These matched and envelope templates will both include the attenuation introduced by the visual field inhomogeneity in contrast sensitivity. Input from more eccentric locations is weighted to have less of an effect on the decision than the input from the fovea, in proportion to the expected signal to noise ratio. Templates which do not feature this variation in weighting over their surface due to the attenuation are referred to as "flat" in this thesis. Where the template is matched to the extent of the stimulus envelope (e.g. is a continuous square or circle, even when the stimulus contains holes), the noise remains in the areas within the stimulus that do not feature signal (Figure 5.5c).

5.2.6 Spatial summation and calculation of the detection threshold

The signals from each location in the stimulus image are combined, either through a linear sum of the pixel values or a max operation over the image. For stochastic models these values are then provided to the decision mechanism. In a simulation of a 2IFC experiment the output from one interval will then be compared against the output from another interval in order to choose the one most likely to contain the stimulus. Repeating this many times with signals of different strengths allows an experiment to be simulated, the results of which give the model prediction (threshold contrast is found by fitting psychometric functions to the simulated data). For analytic models the predictions are derived by calculating the signal-to-noise ratio at the decision stage (see Section 5.3).

5.3 Analytic predictions for models involving linear summation

5.3.1 Building the summation models

This section will use the component stages outlined above to develop four analytic summation models. In each case the detection thresholds predicted by the models can be found by a single equation operating over the attenuated, filtered, and transduced stimulus image $\mathbf{S}_{\text{rect}}$. Rearranging and simplifying this equation provides a prediction of how threshold should vary with stimulus area if the stimulus has a contrast profile which is effectively "flat". In each case the models predict a relationship of the form $c_{\text{thresh}} \propto \frac{1}{\sqrt[a]{n}}$, which corresponds to a straight line with a gradient of $-a$ on log-log axes.

5.3.2 Linear summation model

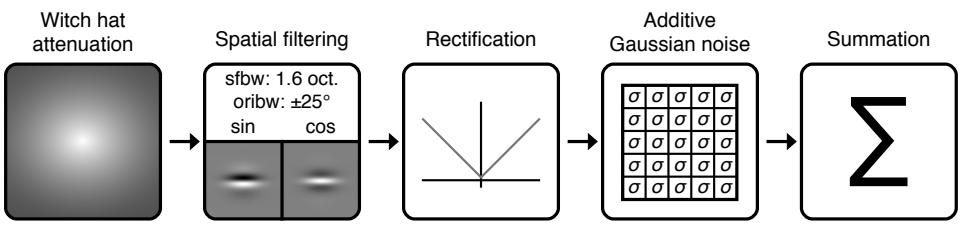

Figure 5.6: Architecture of the linear summation model.

To find the SNR at the decision mechanism for a model involving the linear sum over the image, the signal level of the detector at each monitored location is added. The combined signal is then divided by the standard deviation of the combined noise from these detectors (calculated by summing the variances). For the linear summation model (see Figure 5.6) this takes the form

$$\text{SNR} = \frac{\sum\limits_{x,y} (s_{x,y})}{\sqrt{\sum\limits_{x,y} (\sigma_{x,y}^2)}}, \tag{5.12}$$

where $s_{x,y}$ and $\sigma_{x,y}$ are the pixel values at those locations in $\mathbf{S}_{\text{rect}}$ and $\mathbf{G}$.

For ease of presentation this 2D matrix of responses is reduced to a single dimension (with each detector being the ith in a list of n) in subsequent equations

$$\text{SNR} = \frac{\sum\limits_{i=1}^{n} (c \cdot s_i)}{\sqrt{\sum\limits_{i=1}^{n} (\sigma_i^2)}}. \tag{5.13}$$

As there is no template being applied to the stimulus, the total level of noise (σ_{tot}) remains constant for any stimulus size. The stimulus contrast is a constant factor for the whole stimulus, so this can be factored out of the summation term. This allows us to simplify to

$$\text{SNR} = \frac{c \cdot \sum_{i=1}^{n} (s_i)}{\sigma_{\text{tot}}}. \tag{5.14}$$

The contrast that gives a signal-to-noise ratio of 1 (which is taken to be the detection threshold) can then be found by solving for c

$$c_{\text{thresh}} = \frac{\sigma_{\text{tot}}}{\sum_{i=1}^{n} (s_i)}, \tag{5.15}$$

this is the general version of the linear summation model. For a constant σ_{tot} of 1 and s_i of s

$$c_{\text{thresh}} = \frac{1}{n \cdot s}, \tag{5.16}$$

so for any s

$$c_{\text{thresh}} \propto \frac{1}{n}. \tag{5.17}$$

This means that for a stimulus with a flat contrast profile at the summation stage, the contrast detection threshold should decrease in proportion to the area of the stimulus. Otherwise, the detection threshold will decrease at a rate that can be calculated from Equation 5.14. This will depend on any variation in contrast across the image, which may arise solely from the visual field inhomogeneity (if a stimulus with a flat contrast profile is presented on the display).

5.3.3 Nonlinear and quadratic summation models

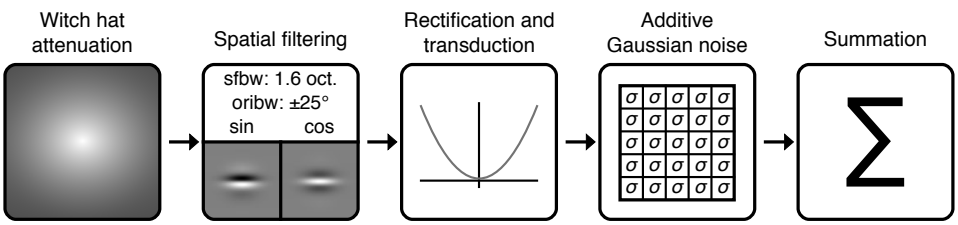

Figure 5.7: Architecture of the nonlinear and quadratic summation models.

For a model similar to that given above in Section 5.3.2, but featuring a nonlinear transducer (see Figure 5.7), the signal from each detector is raised to a power (m)

$$\text{SNR} = \frac{\sum_{i=1}^{n} (c \cdot s_i)^m}{\sqrt{\sum_{i=1}^{n} (\sigma_i^2)}}. \tag{5.18}$$

The contrast term (c^m) can be factored out as it is constant for each location in the image, and as there is no template the noise can be treated as constant for any stimulus size

$$\text{SNR} = \frac{c^m \cdot \sum_{i=1}^{n} (s_i)^m}{\sigma_{\text{tot}}}. \tag{5.19}$$

Solving for threshold contrast at an SNR of 1 gives the general nonlinear summation model

$$c_{\text{thresh}} = \sqrt[m]{\frac{\sigma_{\text{tot}}}{\sum_{i=1}^{n} (s_i)^m}}. \tag{5.20}$$

For a constant σ_{tot} of 1 and s_i of s

$$c_{\text{thresh}} = \sqrt[m]{\frac{1}{n \cdot s^m}}, \tag{5.21}$$

so for any s

$$c_{\text{thresh}} \propto \frac{1}{\sqrt[m]{n}}. \tag{5.22}$$

So in this model the detection threshold of a stimulus with a flat contrast profile at the summation stage should decrease in proportion to the mth root of the area of the stimulus.

For an m of 2, Equation 5.20 is the general version of the quadratic summation model

$$c_{\text{thresh}} = \sqrt{\frac{\sigma_{\text{tot}}}{\sum_{i=1}^{n} (s_i)^2}}. \tag{5.23}$$

In this model, for a constant σ_{tot} of 1 and any fixed s

$$c_{\text{thresh}} \propto \frac{1}{\sqrt{n}}. \tag{5.24}$$

5.3.4 Template and ideal summation models

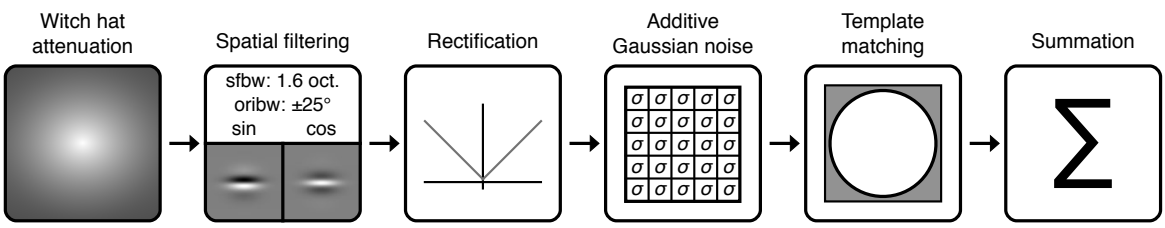

Figure 5.8: Architecture of the template and ideal summation models.

For a model similar to that given above in Section 5.3.2, but featuring a template stage where each detector (represented by a pixel in the filtered and transduced image) is multiplied by its weighting t_i (see Figure 5.8)

$$\text{SNR} = \frac{\sum_{i=1}^{n} (c \cdot s_i \cdot t_i)}{\sqrt{\sum_{i=1}^{n} (\sigma_i \cdot t_i)^2}}. \tag{5.25}$$

The contrast term c can be factored out. The noise cannot be treated as constant because it varies with the size of the stimulus (as it is weighted by the template)

$$\text{SNR} = \frac{c \cdot \sum_{i=1}^{n} (s_i \cdot t_i)}{\sqrt{\sum_{i=1}^{n} (\sigma_i \cdot t_i)^2}}. \tag{5.26}$$

Solving for threshold contrast at an SNR of 1 gives

$$c_{\text{thresh}} = \frac{\sqrt{\sum_{i=1}^{n} (\sigma_i \cdot t_i)^2}}{\sum_{i=1}^{n} (s_i \cdot t_i)}. \tag{5.27}$$

For a constant σ_i of 1 and s_i of s

$$c_{\text{thresh}} = \frac{\sqrt{\sum_{i=1}^{n} (t_i)^2}}{s \cdot \sum_{i=1}^{n} (t_i)}. \tag{5.28}$$

For the case where the template is matched to the stimulus (ideal summation), $t_i = s_i$ so Equation 5.27 simplifies to

$$c_{\text{thresh}} = \frac{\sqrt{\sum_{i=1}^{n} (\sigma_i \cdot s_i)^2}}{\sum_{i=1}^{n} (s_i)^2}, \tag{5.29}$$

this is the general version of the ideal summation* model. For a constant σ_i of 1 and s_i of s

$$c_{\text{thresh}} = \frac{1}{\sqrt{n} \cdot s}, \tag{5.30}$$

so for any s

$$c_{\text{thresh}} \propto \frac{1}{\sqrt{n}}. \tag{5.31}$$

So in this model the detection threshold of a stimulus with a flat contrast profile at the sum-

*A theoretical ideal detector does not feature internal noise. This model is "ideal" *only* given that it is necessary for the modelled system to contain performance-limiting internal noise.

mation stage should decrease in proportion to the square root of the area of the stimulus.

5.3.5 Combination and noisy energy models

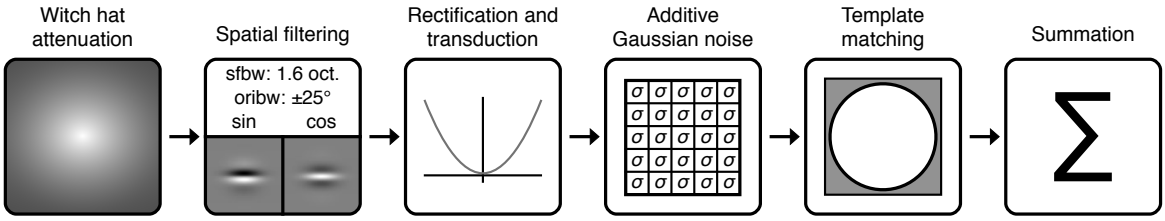

Figure 5.9: Architecture of the combination and noisy energy models.

These models feature both nonlinear transduction and template stages (see Figure 5.9), making their derivation a combination of that presented in Sections 5.3.3 and 5.3.4

$$\text{SNR} = \frac{\sum_{i=1}^{n} \left[(c \cdot s_i)^m \cdot t_i\right]}{\sqrt{\sum_{i=1}^{n} (\sigma_i \cdot t_i)^2}}. \tag{5.32}$$

The contrast term c^m can be factored out. The noise cannot be treated as constant because it varies with the size of the stimulus (as it is weighted by the template)

$$\text{SNR} = \frac{c^m \cdot \sum_{i=1}^{n} (s_i^m \cdot t_i)}{\sqrt{\sum_{i=1}^{n} (\sigma_i^2 \cdot t_i^2)}}. \tag{5.33}$$

Solving for threshold contrast gives

$$c_{\text{thresh}} = \sqrt[m]{\frac{\sqrt{\sum_{i=1}^{n} (\sigma_i^2 \cdot t_i^2)}}{\sum_{i=1}^{n} (s_i^m \cdot t_i)}}, \tag{5.34}$$

which is the general form of the combination model.

For a template model involving transduction, two kinds of matched template can be considered. In the first the template is equal to the stimulus before transduction ($t_i = s_i$), making the threshold contrast

$$c_{\text{thresh}} = \sqrt[m]{\frac{\sqrt{\sum_{i=1}^{n} (\sigma_i^2 \cdot s_i^2)}}{\sum_{i=1}^{n} [s_i^{(m+1)}]}}, \tag{5.35}$$

whereas in the second case the template is equal to the stimulus after transduction ($t_i = s_i^m$). This is the ideal template strategy, making the threshold contrast

$$c_{\text{thresh}} = \sqrt[m]{\frac{\sqrt{\sum_{i=1}^{n} [\sigma_i^2 \cdot s_i^{(2m)}]}}{\sum_{i=1}^{n} [s_i^{(2m)}]}}. \tag{5.36}$$

For a constant σ_i of 1 and s_i of s, Equations 5.35 and 5.36 both simplify to

$$c_{\text{thresh}} = \frac{1}{\sqrt[2m]{n} \cdot s}, \tag{5.37}$$

so for any s

$$c_{\text{thresh}} \propto \frac{1}{\sqrt[2m]{n}}, \tag{5.38}$$

indicating that in this model the threshold for a stimulus with a flat contrast profile at the summation stage should decrease in proportion to the $2m^{\text{th}}$ root of the area of the stimulus.

For an m of 2, Equation 5.36 is the noisy energy model

$$c_{\text{thresh}} = \sqrt{\frac{\sqrt{\sum_{i=1}^{n} [\sigma_i^2 \cdot s_i^4]}}{\sum_{i=1}^{n} [s_i^4]}}, \tag{5.39}$$

a stochastic version of this model was used to account for the area summation data reported in Meese and Summers (2012). For a constant σ_i of 1 and s_i of s

$$c_{\text{thresh}} = \frac{1}{\sqrt[4]{n} \cdot s}, \tag{5.40}$$

so for any s

$$c_{\text{thresh}} \propto \frac{1}{\sqrt[4]{n}}. \tag{5.41}$$

5.4 Analytic approximations

5.4.1 Models without linear summation stages

For model architectures that do not add local outputs in a higher level mechanism but instead achieve improvements in performance with increasing stimulus area through probability summation, the predictions depend on whether the system is assumed to be operating under high threshold theory (HTT, see Section 2.5.5) or signal detection theory (SDT, see Section 2.5.6).

Both versions of the probability summation model are presented here as analytic approximations performed using Minkowski summation

$$\mathsf{f}(\mathbf{X}) = \left(\sum_{i=1}^{n} |x_i|^m \right)^{\frac{1}{m}}. \tag{5.42}$$

Minkowski summation is frequently used in models of signal combination due the flexibility and simplicity with which it can model different behaviours as the exponent (m) changes (Quick, 1974; To, Lovell, Troscianko, & Tolhurst, 2008; Mullen, Beaudot, & Ivanov, 2011; To, Baddeley, Troscianko, & Tolhurst, 2011). For both forms of the probability summation model, Minkowski summation is conducted over the attenuated, filtered, and rectified image $\mathbf{S}_{\text{rect}}$.

5.4.2 HTT probability summation model

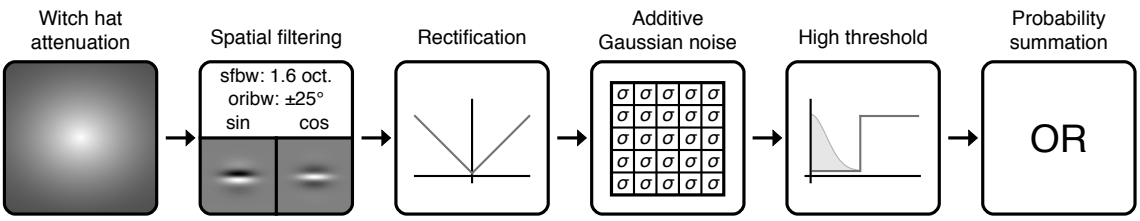

Figure 5.10: Architecture of the probability summation model under HTT.

For the HTT case (see Figure 5.10), the response from each location in the image is perturbed by independent noise. Following this there is a threshold set sufficiently high that it is never (or very rarely) exceeded by the noise alone. If the threshold is exceeded then that detector enters the "detect" state. The system as a whole detects the stimulus if the response in any detector exceeds the threshold. The probability of the detection of the stimulus $P(\text{detect})$ can therefore be derived by combining the probabilities from each individual detector $P_i(s_i)$

$$P(\text{detect}) = 1 - \prod_{i=1}^{n} [1 - P_i(s_i)]. \tag{5.43}$$

The sensitivity (S, the reciprocal of the threshold contrast c_{thresh}) of such a system is approximated by Minkowski summation (Equation 5.42) over the detector outputs (s_i) where $m = \beta$, the slope parameter of a Weibull psychometric function fitted to the data (Quick, 1974; Robson & Graham, 1981). The detection threshold is therefore

$$c_{\text{thresh}} = \left(\sum_{i=1}^{n} s_i^m \right)^{-\frac{1}{m}}. \tag{5.44}$$

5.4.3 SDT probability summation model

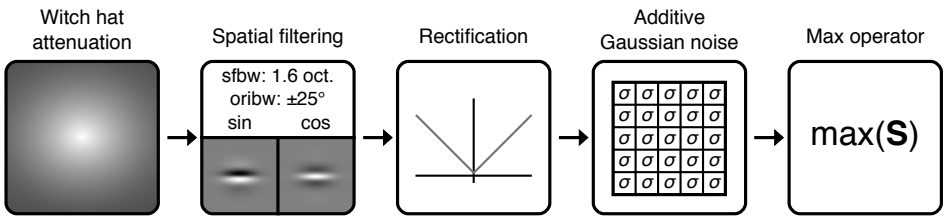

Figure 5.11: Architecture of the probability summation model under SDT.

A significant shortcoming of the original formulation of probability summation is that it was rooted in high threshold theory, the implications of which are at odds with empirical findings (see Section 2.4.2). Probability summation has been reformulated under SDT (Pelli, 1985; Tyler & Chen, 2000) where the high threshold assumption is dropped, and the combination of the detection probabilities of each channel is replaced by a max operator that passes the signal level from the most activated detector. Under the conditions of high uncertainty (see Section 2.3.4) about which channels contain task-relevant signal proposed by Pelli (1985), the effect of summation on threshold is approximated by Equation 5.44 using $m = \beta$ for the values of β typically reported from empirical results ($\beta \approx 4$).

The account of probability summation under SDT has been developed further in a study of the behaviour of stochastic models conducted by Tyler and Chen (2000), who also find that a good approximation is provided by Minkowski summation. In agreement with Pelli (1985), the best fit is found when the exponent m is fixed at 4

$$c_{\text{thresh}} = \left(\sum_{i=1}^{n} s_i^4 \right)^{-\frac{1}{4}} . \tag{5.45}$$

As this equation provides a good approximation to the summation behaviour of a system operating under probability summation given certain conditions are met (a linear transducer and additive Gaussian noise), references made elsewhere in this thesis to "the probability summation model" refer to the predictions made by this equation unless otherwise indicated.

5.5 Conclusions

5.5.1 Summary of model predictions

The models put forward here all predict that log contrast detection threshold should decline in proportion to the log of the stimulus area. The slope of that decline can be used to discriminate

between some models (the linear, matched template, and quadratic summation models all predict shallower summation slopes), but several share the same fourth-root prediction (slope of $-\frac{1}{4}$). This is also the summation slope that is most frequently found in psychophysical experiments. The models predicting this slope are the noisy energy model, the SDT probability summation model, and the HTT probability summation model when $\beta = 4$ (values in this range are frequently found in psychophysical experiments). The HTT probability summation model can be rejected however due to the lack of support for the HTT assumption (see Section 2.4.2).

The two remaining fourth-root models (noisy energy and SDT probability summation), despite making identical predictions for how thresholds should decline as a stimulus increases in size, *do* predict different psychometric function slopes. Meese and Summers (2012) performed a summation experiment where psychometric function slopes were included in the analysis and found results incompatible with SDT probability summation, but in accordance with the predictions made by the noisy energy model. Another way to distinguish between these two models is by designing an experiment in such a way as to bypass the template stage in the noisy energy model. The experiments performed by Meese (2010) with "Battenberg" stimuli (see Section 3.4.6) found $\sqrt{2}$ summation consistent with the quadratic summation model. One explanation for this would be that the noisy energy model is unable to match a template to the Battenberg stimuli, and defaults to summing over the entire display.

CHAPTER 6

Area summation with witch hat compensation

6.1 Motivation and summary

Previous studies investigating the summation of contrast over area in the central visual field have necessarily confounded stimulus extent with the effects of the visual field inhomogeneity in contrast sensitivity. Here, the inverse of the attenuation surface measured in Chapter 4 is applied to stimuli in order to equate the effective contrast across the visual field. The results from this study reveal for the first time that the summation of contrast over the central visual field is both spatially extensive (occurring over at least 33 cycles) and governed by a single fourth-root summation rule. A summation slope of this gradient is consistent with either probability summation or a noisy energy model.

6.2 Introduction

6.2.1 Area summation of low-contrast gratings to threshold

The contrast detection threshold for grating stimuli decreases as a function of the stimulus area. A general review of the results from previous area summation experiments is provided in Section 2.6.1 of the Literature Review. In brief, these studies have found thresholds to decline steeply initially (approximately proportional to the stimulus area for very small stimuli) followed by a period of fourth-root summation (thresholds decline proportional to the fourth-root of the stimulus area), before shallowing off for larger stimulus sizes.

One reason for the summation slopes to flatten for larger stimulus sizes is because the log contrast sensitivity of the human visual system declines as a bilinear function of eccentricity

(see Chapter 4; Baldwin, Meese, & Baker, 2012). The effective contrast of a stimulus therefore decreases as it is moved further from the fixation point. This introduces a confound in area summation experiments, as increments in the size of a stimulus are typically presented at increasingly eccentric locations. Some previous studies have avoided this by either presenting stimuli across regions of the periphery where contrast sensitivity is relatively constant (see Chapter 7; Robson & Graham, 1981; Mayer & Tyler, 1986; Manahilov & Simpson, 2001; Meese & Hess, 2007), or by using "Swiss cheese" (Meese & Summers, 2007, 2009; Baker & Meese, 2011; Meese & Baker, 2011) or "Battenberg" (Meese, 2010) stimuli that allow for the signal area of a stimulus to be varied whilst keeping its extent constant.

The degree of summation found in previous studies that counteracted the effects of the inhomogeneity is still less than that expected from a linear system (which would predict that detection thresholds should decrease in proportion to the stimulus area). Possible nonlinearities in the visual system that would decrease the rate at which the thresholds decline include nonlinear transduction of the input signal (e.g. an energy model; see Manahilov et al., 2001), changes in the amount of integrated noise with increasing stimulus size (e.g. a template model; see Watson & Ahumada, 2005), and nonlinear summation operators (e.g. the max operator; see Tyler & Chen, 2000). Robson and Graham (1981) explained the fourth-root summation found in their study as being due to probability summation (see Section 2.5.5) over independent noisy detectors, which can be approximated by Minkowski summation (see Section 5.4.2; Quick, 1974). This probability summation account (and updated variants, see Pelli, 1985; Tyler & Chen, 2000) has found wide acceptance and is frequently used as a "back-pocket" model of how signals are combined to achieve detection threshold in several domains (e.g. To et al., 2008; Mullen et al., 2011; To et al., 2011).

More recent studies have found levels of summation and psychometric function slopes incompatible with the probability summation account (Manahilov et al., 2001; Meese & Hess, 2007; Meese & Summers, 2007, 2009; Meese, 2010; Meese & Summers, 2012). The results from these studies instead favour a model that combines nonlinear transduction of local filter responses (with an exponent of 2) with semi-flexible task-dependent pooling of signal and noise (template matching with a limited range of possible templates), followed by a linear sum of the pooled responses. This model predicts fourth-root summation for growing grating stimuli (similar to probability summation), but quadratic summation when the template stage fails to match the stimulus. The matched template strategy is thought to fail for distinguishing the Swiss cheese and Battenberg stimuli from ordinary grating stimuli, and it may also do so for stimuli presented in the periphery (though see Chapter 7).

6.2.2 This study

What has not been possible previously however is to investigate area summation across the central visual field without the effect of the visual field inhomogeneity. Stimuli which are pre-

sented in the periphery to counteract this effect can only subtend an area of limited size before they encroach on regions of the visual field with significantly different contrast sensitivity. A method that avoided the effects of the inhomogeneity could reveal the extent that the observer is capable of pooling over. Howell and Hess (1978) found little change in threshold with increasing area for stimuli more than 20 cycles wide, suggesting this was a limit for the pooling of threshold contrast. In a study that used both square and circular gratings, Rovamo, Luntinen, and Näsänen (1993) found a "critical area" beyond which summation did not lower detection thresholds. The size of the critical area depended on the spatial frequency of the grating For 4 c/deg stimuli (as used in this study) performance asymptoted at 65 square cycles, which is equivalent to a diameter of 9 cycles for circular gratings. Another previous study that used Swiss cheese stimuli with a range of modulator frequencies proposed that linear summation of transduced signal occurs over circular 12 - 45 cycle diameter regions, followed by probability summation between those regions (Baker & Meese, 2011).

In this study, I have used my detailed "witch hat" account of the variation in contrast sensitivity over the visual field (Chapter 4) to transform stimuli of the sort used in previous studies (gratings and Swiss cheese modulated gratings) to be equally detectable across their area. With the effects of the inhomogeneity factored out, the prediction of all of the models presented in Chapter 5 is that the log contrast sensitivity should decline linearly with the log of the stimulus area. The slope of the decline indicates which behaviour the human visual system implements, and any shallowing or flattening of that slope would show a limit on the summation of contrast over area. Witch hat attenuation surfaces were also incorporated into the computer modelling. These stimuli and modelling methods were used to measure the summation behaviour of the visual system without the confounding effects of the visual field inhomogeneity. The results from this study show full summation of contrast over 33 grating cycles, suggesting that the shallowing of the measured summation curves in previous studies (within this range, at least) is due to the inhomogeneity rather than any changes in summation behaviour with increasing stimulus extent.

6.3 Methods

6.3.1 Equipment

Stimuli were stored in a CRS ViSaGe and presented on a gamma-corrected CRT monitor (Eizo Flexscan T68). 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 1.19 metres, having a resolution of 48 pixels per degree of visual angle at this distance (12 pixels/cycle for the 4 c/deg stimuli used here).

	Stimulus diameter (cycles)							
Sizes	1.3	2.3	4.3	8.3	11.7	19.0	26.0	33.0
	Signal area (cycles^2) above half-magnitude							
Grating	1.4	4.3	14.8	54.5	106.9	283.5	530.9	855.3
Swiss cheese	Black check			30.5	49.7	143.3	252.4	409.0
	White check			20.8	48.7	126.6	260.5	422.9

Table 6.1: Stimulus diameters and the signal areas where the stimulus is above half-magnitude for the grating and "Swiss cheese" stimuli used in this study.

6.3.2 Stimuli

Two types of stimuli were used: circular 4 c/deg sine-phase horizontal gratings, and Swiss cheese modulated versions of those same gratings (see General Methods Sections 3.4.1 and 3.4.3 for general information about these stimuli). Stimuli were windowed by raised-cosine envelopes with a bandwidth (λ) of 12 pixels. The Swiss cheese modulations had a spatial frequency of 0.8 c/deg and were performed in both cosine ("white" check) and anticosine ("black" check) phase. Eight stimulus sizes were used for the gratings, with five of those also being used for the Swiss cheeses. The diameters and signal areas of the stimuli are shown in Table 6.1.

Stimuli were presented both with flat contrast profiles (uncompensated) and with witch hat compensation for the visual field inhomogeneity (see Figure 6.1). The compensated stimuli were multiplied by the inverse of the attenuation surface measured for each observer in Chapter 4, counteracting the effects of the inhomogeneity and therefore giving them an effectively flat contrast profile at the summation stage of the visual system. The nominal contrast of the witch hat compensated stimuli is their contrast at fixation.

Observers fixated centrally. A quad of fixation points (black 2x2 pixel squares) surrounded the stimuli. Observers were able to use the location of the fixation points to infer the stimulus size and the location of the centre of the display. The stimulus duration was 100 ms.

6.3.3 Observers

Data were collected from three observers: ASB, DHB and TSM. The observers were 22, 28 and 46 years old respectively and all three were experienced psychophysical observers. Optical correction appropriate for the viewing distances tested was worn when required. All experiments were performed binocularly with natural pupils.

6.3.4 Procedures

Thresholds were measured using a two-interval forced-choice (2IFC) three-down, one-up staircase procedure (see General Methods chapter, Section 3.5). Feedback on correctness of re-

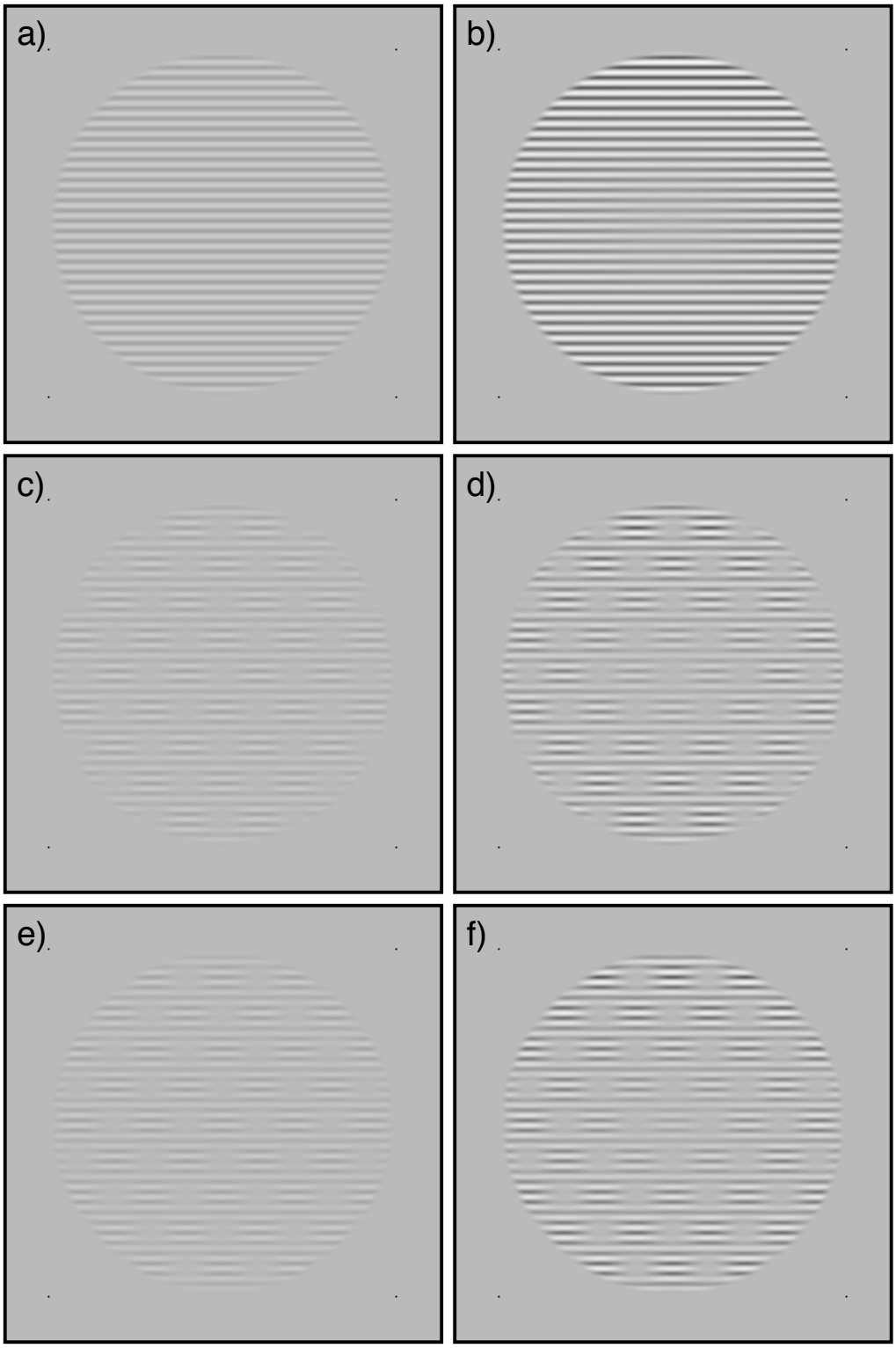

Figure 6.1: Example uncompensated and witch hat compensated grating (a-b), “black” check “Swiss cheese” (c-d), and “white” check Swiss cheese (e-f) stimuli used in this study. Those shown are the largest of the stimuli used in this study (33 cycle diameter).

sponse was provided. Each condition was repeated four times by each observer, except for the Swiss cheese data for observer TSM where data was only gathered for two repetitions. The standard error bars shown in Figure 6.3 indicate that these data were sufficient to constrain the estimate of the threshold for this observer. The contrast detection thresholds for each repetition were calculated using a Weibull fit to the data in Palamedes (see General Methods chapter, Section 3.6), allowing the mean and standard error to be calculated across repetitions. Average psychometric function slopes (Weibull β) were also calculated, from per-repetition slopes constrained to have a maximum value of $\beta = 10$. This constraint was necessary as the error surface for the slope parameter of a fitted psychometric function becomes very shallow when the data are sampled from a function with a steep slope.

6.4 Results

6.4.1 Gratings

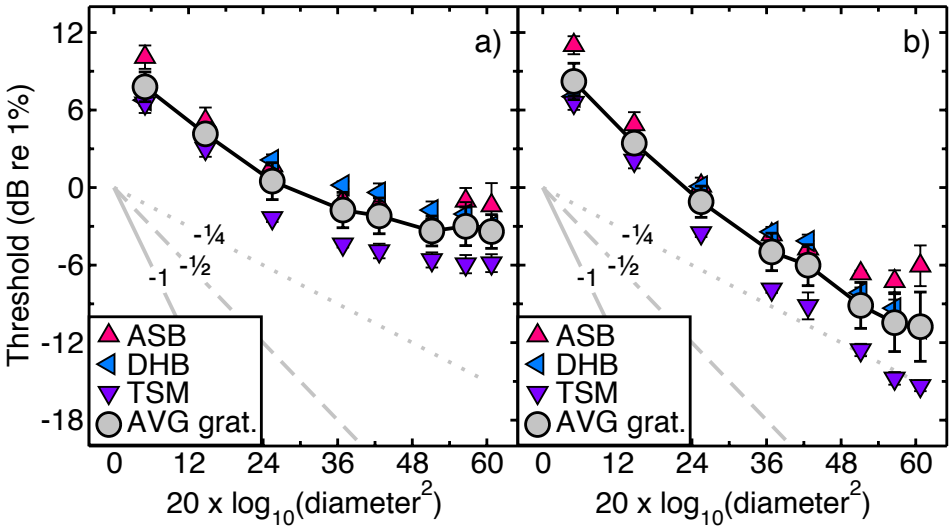

Figure 6.2: Contrast detection thresholds for the grating stimuli from ASB, DHB & TSM, and the average of those data. Log-threshold (in dB) is plotted as a function of the log of the squared stimulus diameter (the square of the diameter multiplied by a factor of $\frac{\pi}{4}$ gives the stimulus area, which would be a shift of -2.1 units on this axis). Panel a) shows the results for the flat stimuli, panel b) shows the results for stimuli that had been multiplied by the inverse of the witch hat attenuation surface. The solid, dashed, and dotted grey lines show slopes of -1, $-\frac{1}{2}$, and $-\frac{1}{4}$ respectively. Error bars show $\pm$ 1 standard error here and in all future graphs.

Threshold contrasts for the detection of the circular grating stimuli are presented in Figure 6.2. Results are shown from the three observers, and the average of those data. The initial steep section of the summation slopes for both the uncompensated (a) and the witch hat compensated (b) thresholds is explained by linear summation within the receptive fields of simple cells

(Meese, 2010). For the uncompensated stimuli (Figure 6.2a), the rate of threshold decline then decreases from fourth-root to more shallow for larger stimulus sizes. This is consistent with a diminishing return in effective signal for larger stimuli due to the effects of the visual field inhomogeneity in contrast sensitivity. For the witch hat compensated stimuli (Figure 6.2b) thresholds declined in proportion to the fourth root of the stimulus area (grey dotted line). Summation slopes of this gradient have previously been explained by probability summation (Robson & Graham, 1981; Tyler & Chen, 2000) or noisy energy (Meese & Summers, 2012) models.

The three observers show some individual differences. For observer ASB thresholds level off for the largest stimulus sizes (19 - 33 cycles), implying a maximum area that can be pooled over to improve performance at detection threshold. The other two observers do not show this effect. For observer TSM thresholds decrease at a slightly steeper rate than that shown by the other two observers. Individual differences such as these have been found previously (e.g. Baker & Meese, 2011), possible explanations for these will be returned to in the Discussion.

6.4.2 Swiss cheese

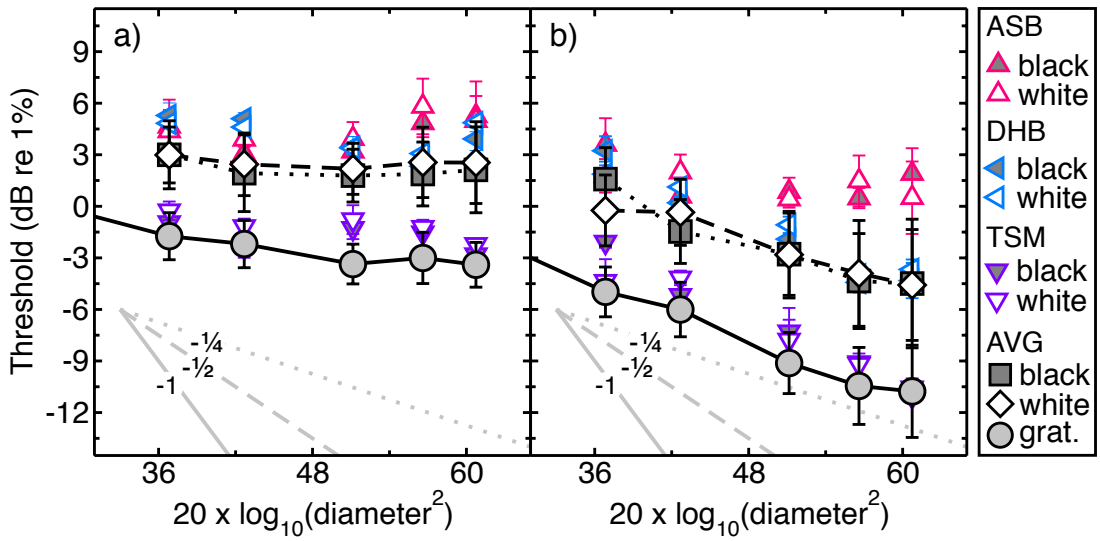

Figure 6.3: Contrast detection thresholds for the "Swiss cheese" stimuli from ASB, DHB & TSM, and the average of those data. Average grating data are also replotted from Figure 6.2. Log-threshold (in dB) is plotted as a function of the log of the squared stimulus diameter (the scales of both axes are different from those in Figure 6.2). Panel a) shows the results for the flat stimuli, panel b) shows the results for stimuli that had been multiplied by the inverse of the witch hat attenuation surface. The solid, dashed, and dotted grey lines show slopes of -1, $-\frac{1}{2}$, and $-\frac{1}{4}$ respectively.

Figure 6.3 shows results from the three observers, and the average of those data, for the detection of the Swiss cheese modulated stimuli. Average thresholds for the grating stimuli are also replotted here from Figure 6.2. Thresholds for Swiss cheese stimuli were 4.5 to 6 dB higher than those for grating stimuli of the same spatial extent (but approximately half the

equivalent signal area due to the contrast modulation, see Table 6.1). Swiss cheese stimuli with the black and white modulator phases had similar detection thresholds. As with the grating stimuli of the same diameter, thresholds do not decrease for non-compensated stimuli (Figure 6.3a) as they increase in size (due to the effects of the visual field inhomogeneity), but thresholds for the witch hat compensated stimuli (Figure 6.3b) decrease parallel to those for gratings (fourth-root summation).

There were some differences between the behaviours for the three observers. For ASB, detection thresholds did not decrease with stimulus size for the larger Swiss cheese stimuli, even when the stimuli were compensated for the visual field inhomogeneity. This is consistent with the effect seen in the grating data for this observer, supporting a maximum contrast integration extent based on stimulus diameter that is within the range tested in this study. For DHB and TSM however, thresholds did decrease with increasing stimulus size.

6.4.3 Psychometric function slopes

The estimated psychometric function slopes from this study may be of interest for comparison against those found in previous studies or for use in models that rely on the slope of the psychometric function as a parameter (e.g. the high threshold theory formulation of probability summation that uses Weibull β as the Minkowski exponent). Consistent with previous studies (Mayer & Tyler, 1986; Meese & Summers, 2012; Wallis, Baker, Meese, & Georgeson, 2013), psychometric function slopes did not vary systematically over the different summation conditions. Across both the grating and Swiss cheese conditions the averaged slopes (Weibull β) of the psychometric functions were 3.4, 4.8, and 4.2 respectively for observers ASB, DHB and TSM. Across all three observers the average slope was 4.0. Previous studies by Robson and Graham (1981), Mayer and Tyler (1986) and Meese and Summers (2012) found average slopes of 3.5, 3.5, and 3.6 respectively. The psychometric function slopes in this study were therefore slightly steeper than those reported previously. Steeper slopes are expected from fits to data collected using adaptive staircase methods (Kaernbach, 2001), and can also result from undersampling of the psychometric function (Wichmann & Hill, 2001; Wallis et al., 2013).

6.5 Modelling

6.5.1 Analytic models of area summation

Eleven area summation models based on those described in Chapter 5 were considered. These were the linear (Section 5.3.2), quadratic (Section 5.3.3), and SDT probability summation (Section 5.4.3) models, and four versions each of the template (Section 5.3.4) and noisy energy (Section 5.3.5) models. These models are listed in Table 6.2. The HTT probability summation

Model variant	Code	Transducer	Template	Sum/max	Equation	Offset (dB)	Grating RMSe (dB)	Swiss cheese RMSe (dB)
Linear	L	Linear	None	Sum	$c_t = \frac{\sigma_{\text{tot}}}{\sum_{i=1}^{n}(s_i)}$	15.60	12.06	10.21
Quadratic	Q	Quadratic	None	Sum	$c_t = \sqrt{\frac{\sigma_{\text{tot}}}{\sum_{i=1}^{n}(s_i)^2}}$	3.78	4.17	3.99
Matched template	T	Linear	Matched	Sum	$c_t = \frac{\sqrt{\sum_{i=1}^{n}(\sigma_i \cdot s_i)^2}}{\sum_{i=1}^{n}(s_i)^2}$	29.78	4.14	3.99
Flat template	FT	Linear	Flat	Sum	$c_t = \frac{\sqrt{\sum_{i=1}^{n}(\sigma_i \cdot t_i)^2}}{\sum_{i=1}^{n}(s_i \cdot t_i)}$	29.69	4.06	3.93
Noisy energy	NE	Quadratic	Matched	Sum	$c_t = \sqrt{\frac{\sqrt{\sum_{i=1}^{n}[\sigma_i^2 \cdot s_i^4]}}{\sum_{i=1}^{n}[s_i^4]}}$	11.53	0.44	1.29
Flat template noisy energy	FTNE	Quadratic	Flat	Sum	$c_t = \sqrt{\frac{\sqrt{\sum_{i=1}^{n}(\sigma_i^2 \cdot t_i^2)}}{\sum_{i=1}^{n}(s_i^2 \cdot t_i)}}$	11.34	0.50	1.21
Extent noisy energy	ENE	Quadratic	Extent	Sum	"	11.53	0.44	0.54
Flat extent noisy energy	FENE	Quadratic	Flat extent	Sum	"	11.34	0.50	0.59
Probability summation	PS	Linear	None	Max	$c_t = \left(\sum_{i=1}^{n} s_i^4\right)^{-\frac{1}{4}}$	-1.13	0.44	1.29

Table 6.2: Models tested in this study, and the codes used in Figures 6.4 and 6.5. Information is provided about each model's transduction and template stages, and whether signals are combined by a linear sum or a max operator. Further explanations of each model are provided in the text. Derivations of the equation for each model are provided in Chapter 5. Also shown are the RMS errors and offset parameters from fitting the model to the grating data, and the RMS error between the "Swiss cheese" data and predictions made for that condition.

model was not implemented as sufficient evidence has been presented in previous studies (see Section 2.4.2; Swets et al., 1961; Green & Swets, 1966e; Nachmias, 1981; Laming, 2013) to reject the high threshold assumption. Regardless, the prediction of the HTT probability summation model would be identical to that for the SDT probability summation model (Tyler & Chen, 2000) in this case, as the average Weibull $\beta = 4$. Each model was fitted to all of the data (from both the flat and the compensated stimuli) using a single free parameter which controlled the global sensitivity (equivalent to adjusting the variance of the noise).

For all models the stimulus was first multiplied by the witch hat attenuation surface to simulate the visual field inhomogeneity in contrast sensitivity, and then filtered with a horizontal sine-phase log-Gabor patch (spatial frequency bandwidth of 1.6 octaves, orientation bandwidth of $\pm 25°$) in analogy to the spatial filtering present at the simple cell level of the visual system. The attenuated and filtered stimulus image was fed into the relevant equation in Table 6.2 to find the threshold contrast (c_t). The pixel values of the attenuated and filtered image were rearranged into a 1D array (s_1 to s_n) of local detector responses, and the standard deviation of the Gaussian noise affecting each detector (σ_i) was assumed to be constant.

The matched template (T) and noisy energy (NE) models both featured templates that were matched to the stimulus exactly ($t_i = s_i$). For the other models featuring template stages these were defined separately. For the flat template (FT) and flat template noisy energy (FTNE) models the template was a version of the stimulus without any variation in contrast over its area from the visual field inhomogeneity or from the witch hat compensation for that inhomogeneity. This models the behaviour of a system that does not represent the expected contrast of the stimulus at each location in its template. For the "extent" noisy energy (ENE) model, the template did not feature the contrast modulation present in the Swiss cheese stimuli (i.e. the template was that of the grating with the same diameter). This represents the behaviour of a system where the range of possible templates is limited to not allow the Swiss cheese contrast modulation to be represented (consistent with the inability to discriminate between gratings and Swiss cheeses at threshold reported by Meese & Summers, 2007). The flat extent noisy energy model (FENE) features both of these constraints on its template stage.

Several of these models make very similar predictions. The probability summation model and the four noisy energy model variants all predict that contrast detection threshold should decline in proportion to the fourth-root of the stimulus area. The matched template and extent template versions of the noisy energy model can be distinguished by their predictions for the Swiss cheese stimuli, with the extent template predicting more summation between the gratings and the Swiss cheeses. The template and quadratic models both predict that thresholds should decrease in proportion to the square-root of the stimulus area. Despite this, all eleven distinct summation models are presented here both for the sake of completeness and to demonstrate the conditions under which they are similar and the conditions under which they are different.

6.5.2 Gratings

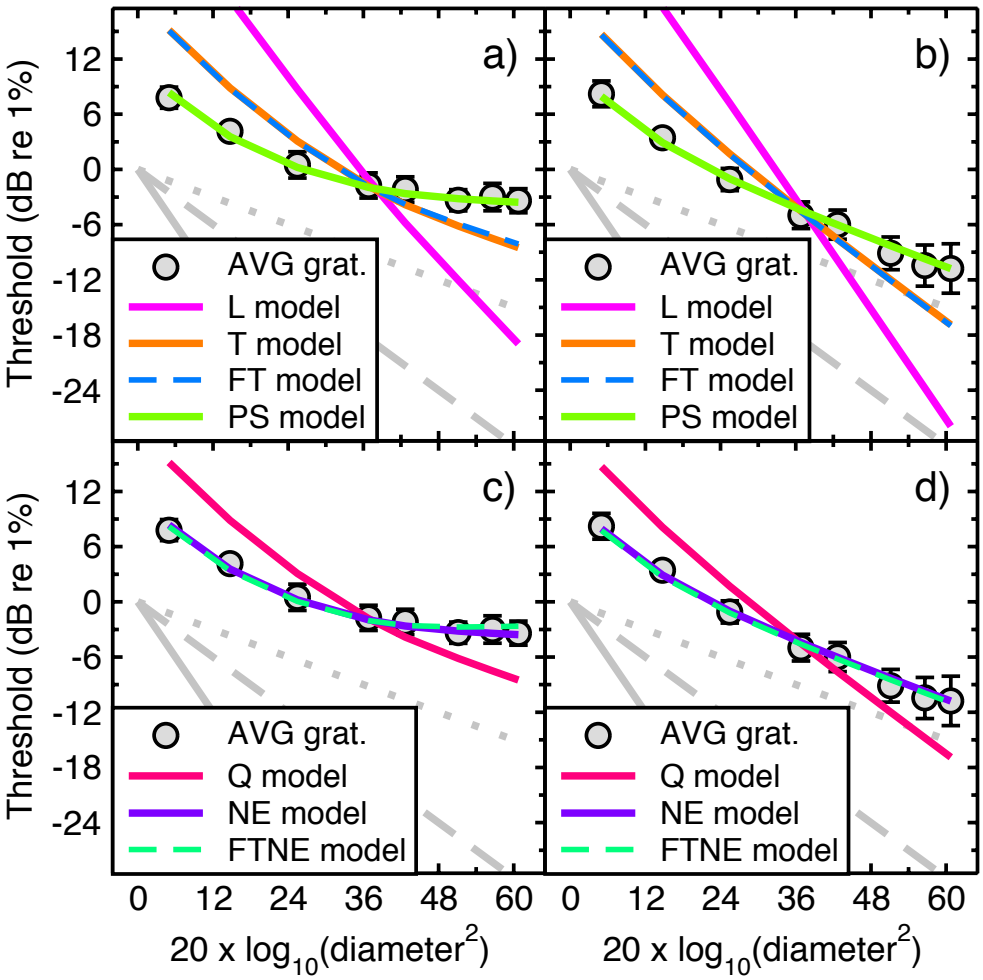

Figure 6.4: Average grating data (replotted from Figure 6.2) fitted by the summation models described in Chapter 5. RMS errors and vertical offset parameters are provided in Table 6.2. Both rows show the same data. Panels a) and c) show the results for the flat stimuli, panels b) and d) show the results for stimuli that had been multiplied by the inverse of the witch hat attenuation surface. In panels a) and b), the data are fitted by the models which involve a linear transducer. These are the linear (L), template (two versions: matched template T and flat template FT), and probability summation (PS) models. In panels c) and d), the data are fitted by the models which feature a nonlinear transducer. These are the quadratic (Q) and noisy energy (two versions: matched template NE and flat template FTNE) models.

Each model was fitted to the averaged grating data using the downhill simplex method in Matlab (see Section 3.7) to minimise the RMS error between the data and the model prediction. The only parameter varied was the global sensitivity of the model (equivalent to varying the standard deviation of the internal noise in these models) which does not effect summaton behaviour, only translating the model prediction curves vertically along the log-threshold axis. Fits were made simultaneously to the thresholds for the non-compensated and witch hat compensated grating stimuli, yielding a single vertical offset (global sensitivity) parameter. Fig-

ure 6.4 shows the model fits to the averaged data, and Table 6.2 provides the RMS errors and vertical offset parameters for those fits. It was not necessary to plot the models with templates matched to the stimulus extent (ENE and FENE), as the predictions for the grating stimuli would be no different from those made by the NE and FTNE models.

For the witch hat compensated stimuli (panel b), the linear summation model (pink curves) predicts a summation slope of -1. This is much steeper than the rate of threshold decline in the results. After the initial steeper decline that results from linear summation within filter elements (simple cell analogues), the quadratic (red curves) and template (orange and blue curves) models both predict summation slopes of approximately $-\frac{1}{2}$, which are also steeper than the decline in threshold seen in the data. The probability summation (green curves) and both versions of the noisy energy model (purple and turquoise curves) predict slopes of approximately $-\frac{1}{4}$, which are consistent with rate of threshold decline in the data. The predictions of the probability summation and matched template noisy energy models are in fact identical (see Section 5.5.1) with both giving an RMS error of 0.44 dB, and the prediction of the noisy energy model with the flat template is very similar (RMSe = 0.50 dB). An extensive previous study investigating the slopes of the psychometric functions and the effects of uncertainty (in addition to summation slopes) by Meese and Summers (2012) favours noisy energy models over probability summation models, however the results presented in that study were not sufficient to constrain any particular version of the template stage (NE or FTNE).

6.5.3 Swiss cheese

The vertical offset parameters from the three best-fitting models presented in Table 6.2 were used to generate predictions for the Swiss cheese stimulus thresholds. These predictions are presented with the average data in Figure 6.5, and the RMS errors between the predictions and the averaged data are presented in right-hand column of Table 6.2. In addition to the model predictions shown in Figure 6.4, two additional model predictions are plotted. These are from the noisy energy models with template matched to the stimulus envelope without the Swiss cheese modulation (i.e. using the envelopes from the grating stimuli of the same size). At a first glance, all of the models presented here provide a good fit to the data.

The model predictions for the probability summation (top row, green and purple curves) and matched template noisy energy (middle row, blue and red curves) models are once again identical, and both provide reasonable fits to the Swiss cheese data (RMSe = 1.29 dB). The flat template noisy energy model (middle row, cyan and orange dashed curves) model is also very similar to that for the PS and NE models and provides only a marginally superior fit (RMSe = 1.21 dB). All three models predict slightly too little summation. Modifying the template stage of the noisy energy model such that it is matched to the stimulus without the Swiss cheese modulation increases the predicted summation for these conditions (essentially removing the effect of the template stage).

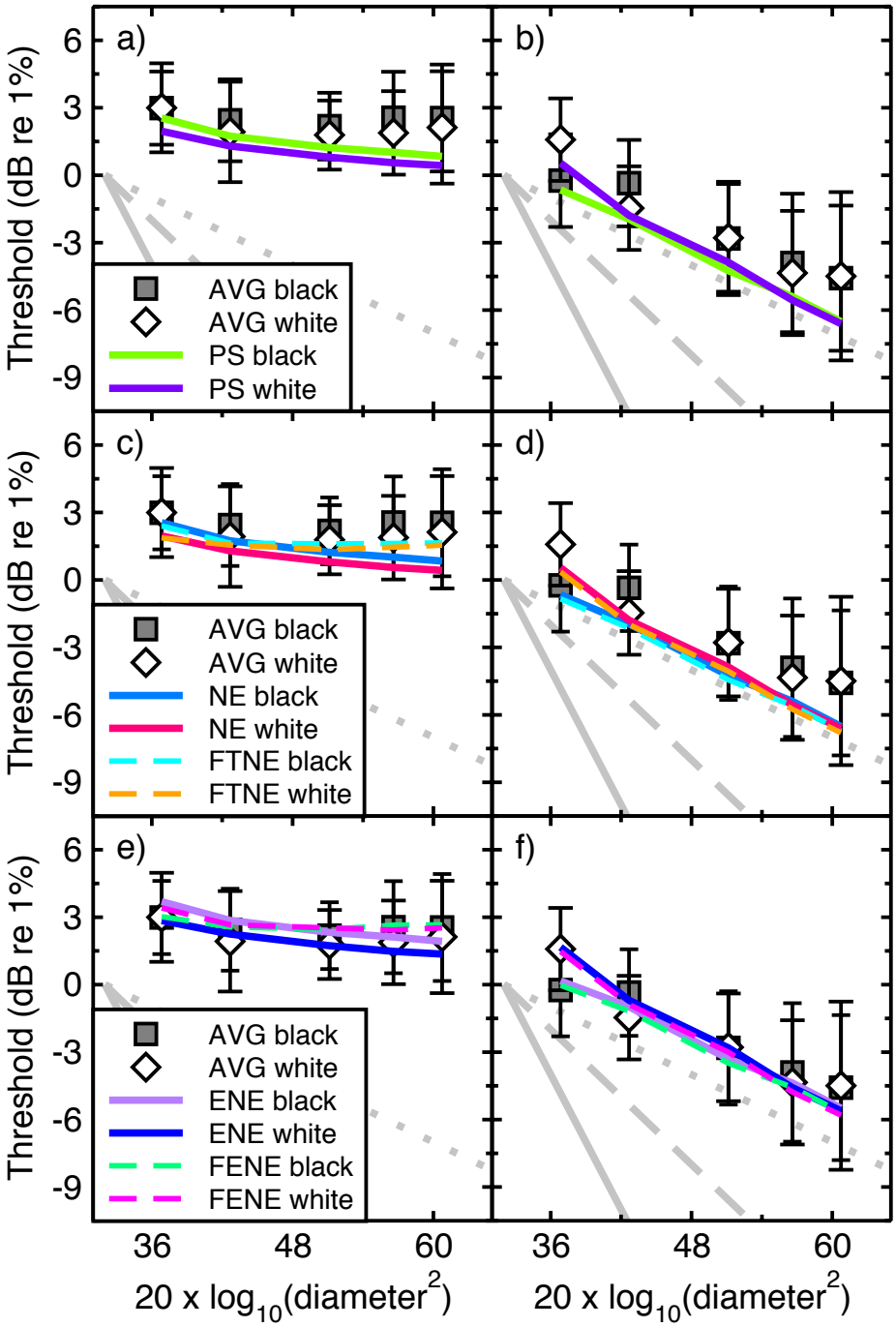

Figure 6.5: Average "Swiss cheese" data (replotted from Figure 6.3) plotted with area summation model predictions based on the fits shown in Figure 6.4. RMS errors are provided in Table 6.2. Results from grating stimuli are omitted from this figure to allow for a finer y-axis scale. All three rows show the same data. Panels a), c), and e) show the results for the flat stimuli. Panels b), d) and f) show the results for stimuli that had been multiplied by the inverse of the witch hat attenuation surface. Panels a) and b) show the probability summation (PS) model. Panels c) and d) show two versions of the noisy energy model with a template matched to the stimulus. Panels e) and f) show another two versions of the noisy energy model with a template matched to the stimulus extent.

The matched template (ENE, bottom row, lilac and blue curves) and flat template (FENE, bottom row, turquoise and pink dashed curves) versions of this model provide a numerically better fit than the other models (RMSe = 0.54 and 0.59 dB respectively), however the predictions from all five models fall within the variability of the data. The results of this study can not therefore be used to reliably distinguish between them. This is partly due to a shortcoming of the Swiss cheese stimulus when used to investigate the summation levels. Because there are large overlapping regions of non-zero contrast in the component black check and white check stimuli, extensive linear summation occurs in those regions regardless of the model architecture. This obscures the detailed summation behaviour of the different model architectures.

6.5.4 Interleaved designs and the matched template

The model predictions made that feature a template matched to the stimulus exactly introduce a complication for studies where stimuli are interleaved. As the observer cannot know which of the interleaved stimuli to expect on each trial they will necessarily be monitoring the template for each possible stimulus on each trial. The observer must then make a decision on a trial-by-trial basis using the template most likely to match the stimulus. Meese and Summers (2012) implemented this in their models by taking the max across the response from the possible stimulus templates on each trial (based on theoretical work by Tyler & Chen, 2000). The most activated template was the one most likely to contain the stimulus, so the decision as to which trial contained the target was made on the basis of its response magnitude.

In this study, the uncompensated and witch hat compensated versions of each stimulus were interleaved. For the flat template models (FT, FTNE and FENE) this does not present an issue as the templates for both versions of the stimuli are identical. For the matched template models (T, NE and ENE) the predicted behaviour is based on an observer who always uses the correct template on each trial. That is to say, the internal noise never causes this observer to select the incorrect template. Under signal detection theory, such an assumption is invalid, however the difference made by incorporating a max-across-templates stage in this model would be so slight that it would make no appreciable difference to the model predictions. The corrected prediction for each model would be intermediate between that made by the matched template version and that made by the flat template version. This issue is explored further in the next chapter (Section 7.5.4).

6.6 Discussion

6.6.1 The integration of contrast over space is extensive

The results and modelling from this study show that the integration of contrast over area is extensive, extending out to at least 33 carrier cycles for two out of the three observers tested

here (DHB and TSM), and for the averaged data. This can be seen in Figure 6.2b, where after the initial steeper slope caused by linear summation within filter elements, summation occurs according to a single rule (fourth-root summation) up to the largest stimulus size tested. Although the noisy energy model (with a template matched to the stimulus extent) provides the best fit to the data presented here, this study cannot reliably distinguish between the predictions of the probability summation and noisy energy models. Based on the evidence from previous studies (e.g. Meese & Summers, 2012), the noisy energy model interpretation of the results is favoured.

The results from this study show summation occured over a greater extent than that reported previously by Howell and Hess (1978) who found the improvement in performance with increasing stimulus area to asymptote at 20 cycles, or by Rovamo et al. (1993) who did not find any summation beyond 9 cycles for 4 c/deg gratings. The results presented here can be reconciled with those of Baker and Meese (2011), whose model featured linear summation of transduced contrast over 12 - 64 cycles followed by probability summation between those integration regions, if the size of the maximum linear integration region in their model were set to at least 33 cycles in diameter.

6.6.2 Individual differences for the extent of area summation

The data for observer ASB show a shallowing of the summation slope for the largest compensated grating and Swiss cheese stimuli (diameters of 19 - 33 cycles). The other two observers show summation over the entire stimulus range tested here (33 cycles). Previous studies have also reported individual differences in the extents over which observers sum contrast (e.g. Baker & Meese, 2011, 2012), however these studies have inferred the summation region size in their modelling rather than being able to demonstrate it directly as is shown here.

The summation slope for observer TSM is steeper than that for either ASB or DHB. This is consistent with previous area summation studies (e.g. Meese & Summers, 2009; Baker & Meese, 2011), where TSM reliably shows more summation than other observers. Previously these data could be explained if the decline in contrast sensitivity with eccentricity was shallower for TSM than for other observers, however the compensation methods used here rule out that explanation. Alternative explanations for the increased summation in this observer would include an increase in receptive field size with eccentricity, a shallower transducer exponent, or an over-estimation of the stimulus extent at the template stage for smaller stimuli.

CHAPTER 7

Area summation across the visual field

7.1 Motivation and summary

Several previous studies have measured the summation of contrast across strips of grating where the bars of the grating are perpendicular to the axis of elongation ("tiger-tail" stimuli). Where these stimuli are presented in the fovea, the results usually show little summation (fourth-root or less) due to the effects of the visual field inhomogeneity in contrast sensitivity. In the periphery however a variety of summation behaviours have been reported ranging from linear (Mayer & Tyler, 1986), to quadratic (Manahilov & Simpson, 2001; Meese & Hess, 2007), to fourth-root (Robson & Graham, 1981; Mayer & Tyler, 1986; Meese & Hess, 2007). An extensive study of summation in the fovea, parafovea, and periphery was carried out in order to try and reconcile these previous conflicting findings. The results of this study show summation behaviour to be identical across the visual field once the inhomogeneity in contrast sensitivity is compensated for. Data collected in the fovea, parafovea, and periphery were well-described by a single noisy energy model.

7.2 Introduction

7.2.1 Spatial summation along strips of grating

As reported in Chapter 6, the threshold contrast for the detection of a stimulus decreases as the stimulus area increases. The shape of this summation slope can reveal properties of the mechanisms in the visual system that combine signals over space. Increasing the size of a stimulus does however necessitate that the increments to the stimulus be presented to locations in the visual field which were not being stimulated by the smaller stimuli. This introduces a confound, as log contrast sensitivity is not homogeneous over the visual field and instead declines in a bilinear fashion with increasing eccentricity from fixation (see Chapter 4).

Publication	Loc.	Stimulus	SF (c/deg)	Carrier	Mod.	A_{min}	A_{maj}	Area (cy^2)
Robson and Graham (1981)	Fov.	Rectangular grating	3-24	H	V	4 cycles	2-64 cycles	20.0-513.4
"	Per.	"	3-12	V	H	"	"	"
Mayer and Tyler (1986)	Par.	Annular grating strip	12	V	H	12 cycles	4-48 cycles	95.7-1583.0
Polat and Norcia (1998)	Fov.	Gabor	3	H & V	H & V	σ = 1 cycle	σ = 1-6 cycles	17.4-104.1
Polat and Tyler (1999)	Fov.	Gabor	4	H & V	V	σ = 1 cycle	σ = 1-6 cycles	17.4-217.4
Manahilov et al. (2001)	Par.	Gabor	2	H & V	H	σ = 0.5 cycles	σ = 1.1-5.6 cycles	2.5-12.6
Foley et al. (2007)	Fov.	Gabor	4	V	H & V	σ = 1 cycle	σ = 0.1-112 cycle	0.2-24.5
Meese and Hess (2007)	Fov.	Rectangular grating	1-4	H & V	H & V	1.7 cycles	1.7-11.7 cycles	3.0-20.3
"	Par.	"	"	"	H & V	"	"	"
This study	Fov.	Battenberg array	4	V	H	0.7-3.7 cycles	1-32 cycles	0.7-118.4
"	Par.	"	"	"	"	"	"	"
"	Per.	"	"	"	"	"	"	"

Table 7.1: Summary of stimulus and experiment details from previous area summation studies that used “tiger-tail” stimuli. Details shown include the stimulus location (foveal, parafoveal, or peripheral), the stimulus type, the spatial frequency, the orientations of both the stimulus carrier and the stimulus envelope, the range of widths for the minor and major axes of the stimulus envelope (full width at half magnitude), and the area of the stimulus calculated from those dimensions. The stimulus sizes from this study are also presented according to this convention here (despite being presented as the area where the contrast was above zero elsewhere in this chapter). The results of all of the previous studies are shown in Figures 7.1, 7.2 and 7.3.

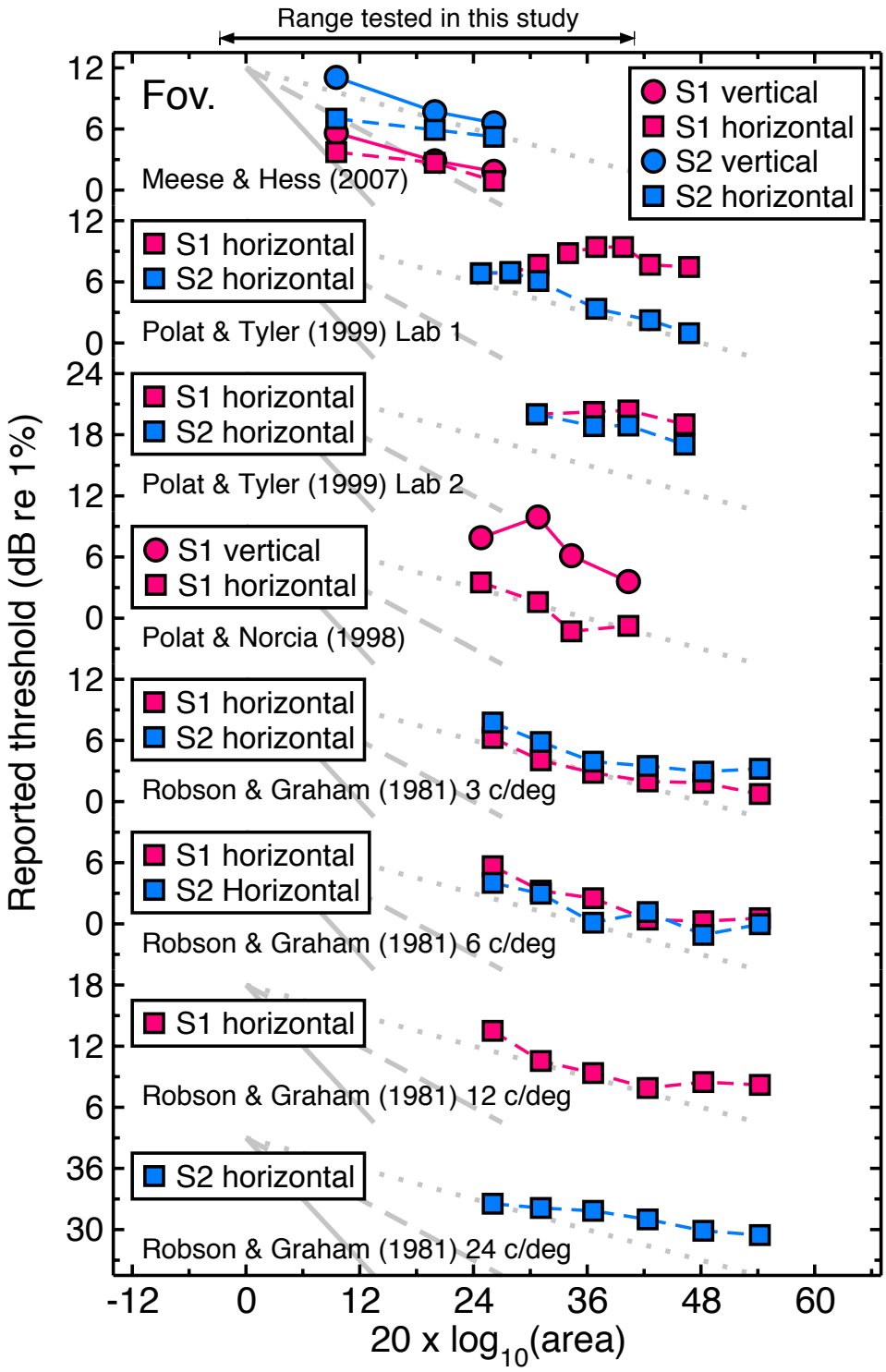

Figure 7.1: Results from previous studies that presented "tiger-tail" stimuli in the fovea (these are listed in Table 7.1). Contrast detection thresholds were extracted from the results presented in those publications and plotted (in dB re 1%) against the log stimulus area. Results are plotted separately for each observer (S1 - S2) in each study, and for whether the orientation of the grating stripes in the stimulus was horizontal or vertical. The solid, dashed, and dotted grey curves show slopes of -1, $-\frac{1}{2}$ and $-\frac{1}{4}$ respectively.

Robson and Graham (1981) avoided the variation in sensitivity over the visual field in their study of area summation by positioning the stimuli in a location where they had measured the contrast sensitivity to be homogeneous (a horizontal strip 42 carrier cycles above the fixation point). Here they found the contrast detection threshold to decrease in proportion to the fourth root of the stimulus area. This was presented as evidence for probability summation under high threshold theory, however this result is also consistent with probability summation under signal detection theory or with a noisy energy model (see Section 6.5).

7.2.2 Summation in the fovea and the periphery

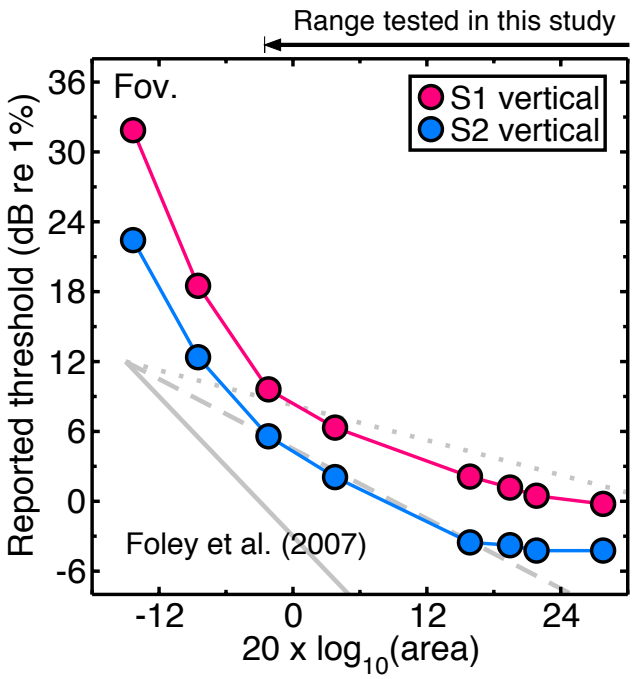

Figure 7.2: Results from Foley et al. (2007), where "tiger-tail" stimuli were presented in the fovea (the details of this study are listed in Table 7.1). Contrast detection thresholds were extracted from the results presented in this publication and plotted (in dB re 1%) against the log stimulus area. Results are plotted separately for each observer (S1 - S2) in the study. The solid, dashed, and dotted grey curves show slopes of -1, $-\frac{1}{2}$ and $-\frac{1}{4}$ respectively.

Several other studies have also investigated area summation along rectangular or oval strips of grating. Where the orientation of the carrier grating is orthogonal to the elongated axis these are sometimes known as "tiger-tail" stimuli. Stimuli with their major axis aligned *along* the grating stripes are sometimes known as "skunk-tail" stimuli. Skunk-tail stimuli typically show more summation than tiger-tail stimuli, however the reported magnitude of this effect has varied between studies (Polat & Tyler, 1999; Foley et al., 2007; Meese & Hess, 2007). The increased summation for skunk-tails has been explained by receptive fields elongated along

the bars of the grating (leading to more within-filter summation, see Section 5.2.2), preferential collinear connections, or mismatched templates (Polat & Tyler, 1999; Meese & Hess, 2007). As these effects would complicate this initial attempt to examine the summation behaviour across the visual field, the scope of this study is restricted to tiger-tail stimuli.

A summary of the methods used in previous tiger-tail summation studies is presented in Table 7.1. The results from all but one of the foveal studies are replotted here in Figure 7.1. In general these show fourth-root or shallower summation slopes as the stimulus grows in size, however they will all be confounded by the visual field inhomogeneity in contrast sensitivity. Figure 7.2 shows the replotted data from Foley et al. (2007). This study tested much smaller stimuli in the fovea (down to 0.2 cycles^2) and found a very steep summation slope. This increased summation for very small stimuli is predicted by within-filter summation if the smallest simulated receptive fields available are larger than the stimuli.

In the periphery a wider range of summation behaviours are seen (Figure 7.3). The larger stimuli of Mayer and Tyler (1986) and some conditions with smaller stimuli in Meese and Hess (2007) replicate the fourth-root summation seen in Robson and Graham (1981). Other conditions in Meese and Hess (2007) and Manahilov et al. (2001) find steeper, quadratic summation. The conditions in Mayer and Tyler (1986) that use smaller stimuli find steep (linear) summation, over a range of sizes where such an effect is not seen in other studies. The variety in the previously reported summation results means that at least one study can be found to support each of a wide range of model architectures, including those that contradict each other (for example, any of the models presented in Chapter 5 would predict either the linear, quadratic, or fourth-root summation slopes listed above).

One interesting feature of the data presented in Figure 7.3 is that studies that found quadratic summation (Meese & Hess, 2007; Manahilov et al., 2001) tested with smaller stimuli than were used in the study that found fourth-root summation (Robson & Graham, 1981). In fact, the steeper summation slope seen for the leftmost points in the 3 c/deg data from Robson and Graham (1981) indicates that the two datasets might each be sampled from the same smooth summation curve (which is initially quadratic before shallowing to fourth-root). The range of stimulus sizes used in this study (indicated by the label above the plot in Figure 7.3) has been chosen to investigate this possibility.

7.2.3 This study

Due to the differences in summation that have previously been identified between the fovea and the periphery, and the apparent contradictions of the results from previous studies (Robson & Graham, 1981; Mayer & Tyler, 1986; Manahilov & Simpson, 2001; Meese & Hess, 2007), an experiment was carried out to investigate spatial summation over strips of grating at three locations in the visual field. The locations chosen were the fovea (0 degrees eccentricity), the

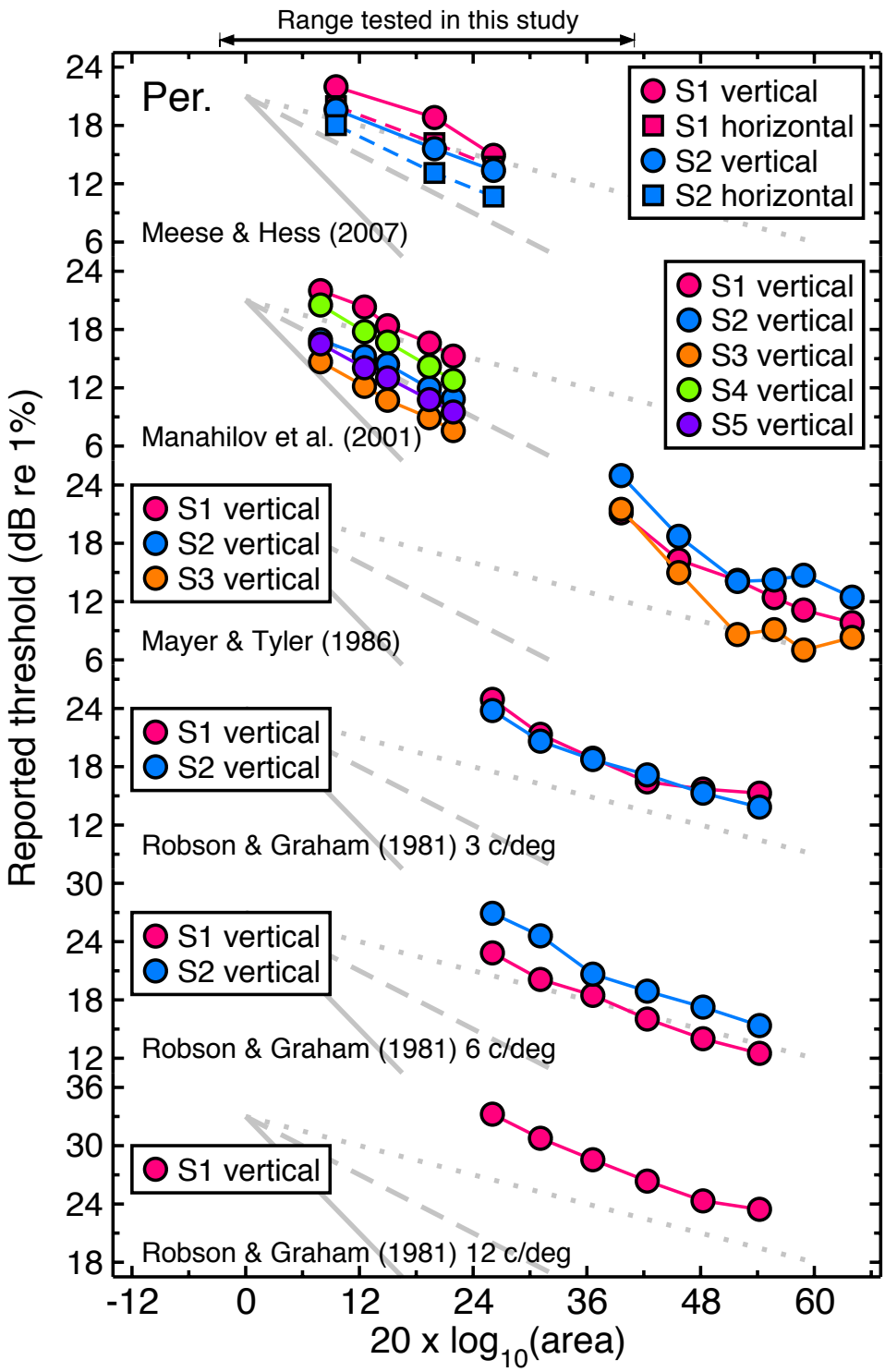

Figure 7.3: Results from previous studies that presented "tiger-tail" stimuli in the periphery (these studies are listed in Table 7.1). Contrast detection thresholds were extracted from the results presented in those publications and plotted (in dB re 1%) against the log stimulus area. Results are plotted separately for each observer (S1 - S5) in each study, and for whether the orientation of the grating stripes in the stimulus was horizontal or vertical. The solid, dashed, and dotted grey curves show slopes of -1, $-\frac{1}{2}$ and $-\frac{1}{4}$ respectively.

parafovea (3 degrees) and the periphery (10.5 degrees). The stimuli were presented both with a constant contrast over area ("flat"), and multiplied by the inverse of the attenuation surface described in Chapter 4 such that they were equally detectable at any location ("witch hat compensated", see Section 4.9.5). The effects of the visual field inhomogeneity therefore no longer confound the investigation of area summation for these stimuli.

This study reveals a summation of contrast over area that is remarkably similar at the three different locations tested, with a fourth-root summation slope consistent with a noisy energy model (Meese & Summers, 2012). In nearly all cases, applying the witch hat compensation equalises the performance across the visual field. The predictions from this single model fitted the data across the stimulus size range tested in this study. Possible reasons for the inconsistent results found previously are considered in the discussion.

7.3 Methods

7.3.1 Equipment

Stimuli were stored in a CRS ViSaGe and presented on a gamma-corrected CRT monitor (Eizo Flexscan T68). 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 1.19 metres. At this distance there were 48 pixels per degree of visual angle, giving the 4 c/deg stimuli used in this study 12 pixels per cycle.

7.3.2 Stimuli

Tiger-tail rectangular grating stimuli were formed from 4 c/deg vertical "Battenberg" elements (see General Methods chapter, Section 3.4). Stimuli were presented with six different widths (1, 2, 4, 8, 16 and 32 cycles) and three different heights (1, 2 and 4 cycles). The areas of the stimuli were calculated as the product of the width and height where the stimulus envelope was above zero (note that this is different from how the areas were calculated in Table 7.1, where the full width at half magnitude was used). These sizes were chosen to cover most of the range tested in previous studies (as demonstrated by the labels above Figures 7.1 to 7.3).

Stimuli were generated both with flat contrast profiles and with witch hat compensation for the visual field inhomogeneity. The compensated stimuli were multiplied by the inverse of the witch's hat attenuation surface measured for each observer in Chapter 4 in order to give them an effectively flat contrast profile following the inhomogeneity. The nominal contrast of the compensated stimuli was their contrast before the surface was applied. This is equivalent to the contrast at fixation (where the attenuation surface was at unity) for the foveal stimuli.

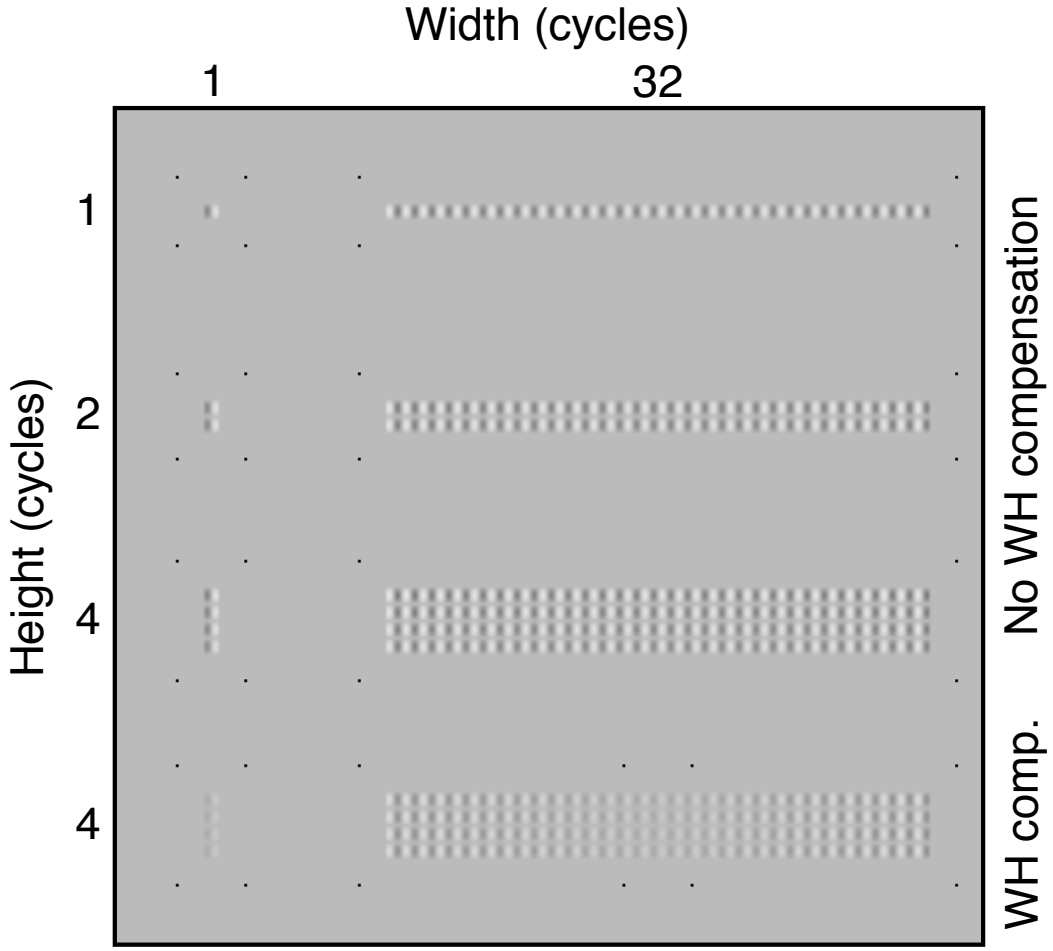

Figure 7.4: Examples of stimuli used in this study. The narrowest (1 cycle) and widest (32 cycles) examples of the non-compensated stimuli of each stimulus height (1, 2 and 4 cycles) are shown. Also shown are the quads of pixels shown around these stimuli that indicated the stimulus extent to the observer. Foveal witch hat compensated stimuli are shown for the the 1x4 and 32x4 stimulus conditions. These stimuli demonstrate the fixation paradigm used in the fovea, where fixation location was inferred from a quad of points (coincident with the stimulus extent quad for the 1x4 condition).

7.3.3 Observers

Data were collected from three observers: ASB, DHB and SAW. The observers were 22, 28 and 46 years old respectively. Observers wore optical correction appropriate for the viewing distances tested when required. All experiments were performed binocularly with natural pupils.

7.3.4 Procedures

Stimuli were presented both centred on fixation and at two eccentric locations (12 and 42 stimulus carrier cycles, equivalent to 3 and 10.5 degrees of visual angle) as shown in Figure 7.5.

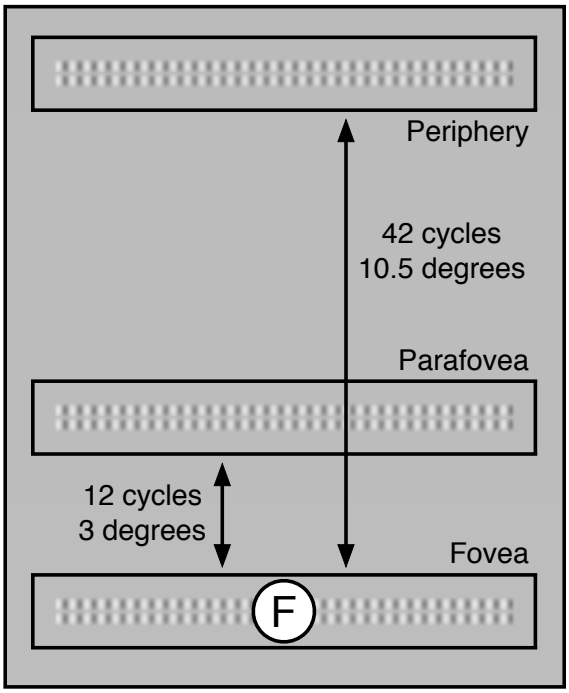

Figure 7.5: Diagram showing the stimulus locations tested in this study (not to scale). The circled "F" marking the fixation location was not present in the study. The fixation location was either inferred from the quad of points around the stimulus (in the fovea) or indicated by a 2 pixel square black dot (parafovea) or red LED (periphery).

These three conditions are labelled as testing the fovea (taken to be the area within 2 degrees of fixation), the parafovea (2 - 10 degrees), and the periphery (beyond 10 degrees) respectively. As the stimuli were extended over space however these labels refer only to the eccentricity of the stimulus centre, and there was some overlap in the eccentricities tested (the outer edges of the largest stimuli were at 4 deg, 5 deg, and 11.2 deg respectively for the foveal, parafoveal, and peripheral conditions).

Observers ASB and DHB were tested at all three locations. Observer SAW was tested only in the fovea and parafovea. Stimulus extent was indicated by a quad of black dots that surrounded the corners of the stimulus. For the foveal and parafoveal conditions these dots were 2 x 2 pixel squares. To ensure visibility in the peripheral condition, the size of these dots was increased to 4 x 4 pixels. Depending on the stimulus location, the intended fixation location was indicated either by an additional quad of dots around that location (for stimuli presented in the fovea), by a single 2 pixel square dot (for stimuli presented in the parafovea), or by a dim red LED positioned at the appropriate distance below the monitor (for stimuli presented in the periphery).

Stimuli were blocked by size and location, but the non-compensated and witch hat compensated conditions were interleaved. Thresholds were measured using a two-interval forced-choice (2IFC) three-down, one-up staircase procedure (see Section 3.5). Feedback on correctness of response was provided. Each condition was repeated four times by each observer. The

contrast detection thresholds for each repetition were calculated using a Weibull fit to the staircase data in Palamedes (see General Methods chapter, Section 3.6), allowing the mean threshold and standard error to be calculated across repetitions.

7.4 Results

7.4.1 Foveal and parafoveal tiger tail summation

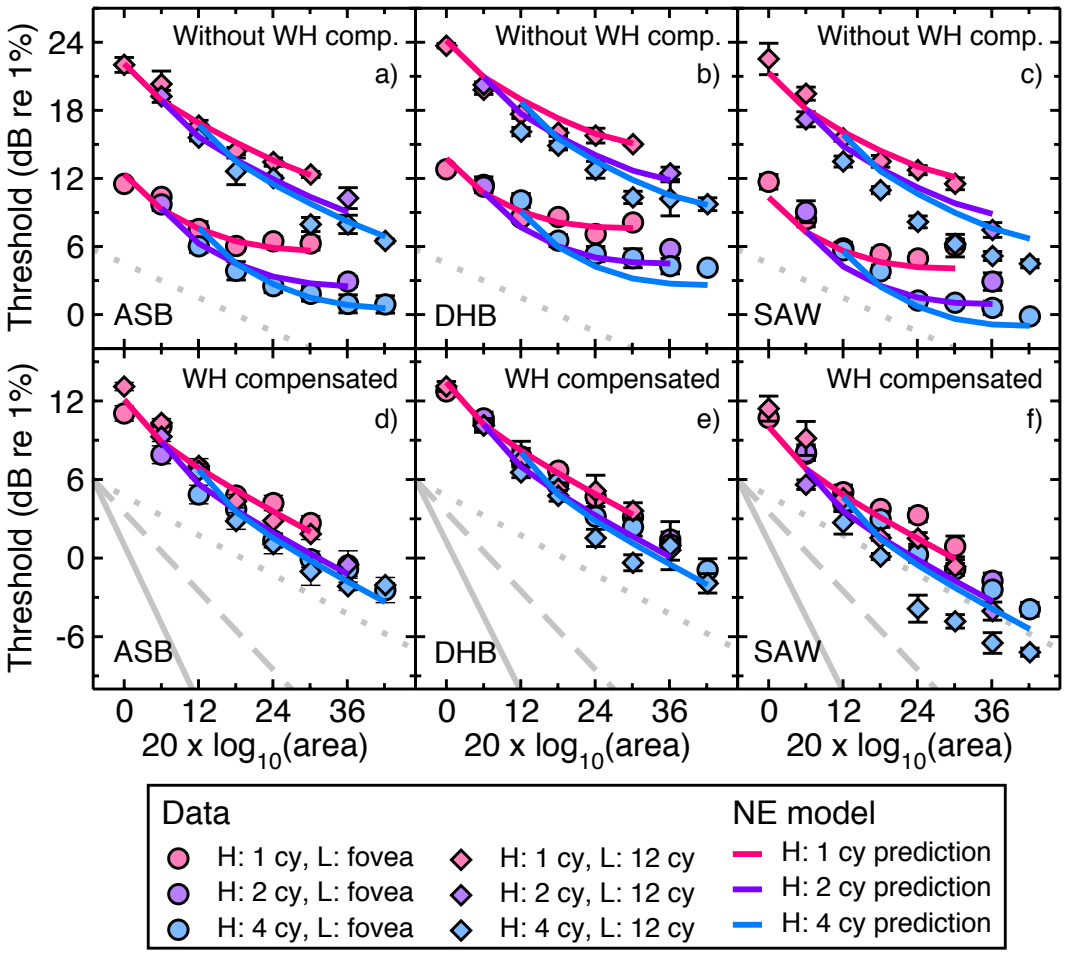

Figure 7.6: Detection thresholds for "tiger-tail" stimuli presented in the fovea and the parafovea (12 cycles eccentricity). Results are shown separately for the three observers. The top three panels (a-c) show thresholds for the condition with the non-compensated stimuli, the bottom three panels (d-f) show thresholds for the witch hat compensated stimuli. The data are fitted by predictions from the noisy energy (NE) model (see Section 7.5). In panels a) to c) the predicted behaviour for the two stimulus locations is shown by the two vertically-offset curves plotted for each stimulus height. In panels d) to f) the model predictions are identical, and so the two curves overlap. The solid, dashed, and dotted grey curves in this and all other figures show slopes of -1, $-\frac{1}{2}$ and $-\frac{1}{4}$ respectively.

The results from the three observers for the foveal and parafoveal condition are shown in Figure 7.6 (fitted by the noisy energy model, see Section 7.5). Results for the periphery are shown in the next section, for convenience of presentation. Thresholds decrease with increasing stimulus area for all conditions. For the non-compensated stimuli (panels a-c), summation in stimuli centred on the fovea (circle symbols) is initially steeper, and then flattens out for larger stimulus sizes due to the visual field inhomogeneity in contrast sensitivity. This is the same result as has been found in the fovea previously (see Figure 7.1).

Applying the witch hat compensation for the inhomogeneity (panels d-f) straightens out the summation slopes to reveal extensive summation of contrast (with a slope of $-\frac{1}{4}$) in the fovea that is not apparent with the non-compensated stimuli (as in Chapter 6). Summation extends across 32 cycles of the 4 c/deg stimulus, equivalent to 8 degrees of visual angle. The witch hat compensation has less of an effect on the summation slopes for stimuli centred on the parafovea (compare diamonds in panels a-c with those in d-f) due to the relative homogeneity of sensitivity in this region. Both stimulus conditions show fourth-root summation over 32 cycles in the parafovea, which is shallower than that found by Manahilov et al. (2001) at around this eccentricity (Figure 7.3).

Applying the witch hat compensation to the stimuli equates thresholds in the fovea and the parafovea. This suggests that i) the attenuation surfaces are accurate, and ii) summation in these locations occurs according to the same rule once the inhomogeneity is factored out. For observers ASB and DHB (panels d-e), the symbols for the foveal and parafoveal conditions (circles and diamonds respectively) overlap for most stimulus sizes. For observer SAW there is some disparity between the results for these two locations. The summation slope for the 4 cycle high stimuli is steeper in the parafovea than in the fovea (panel f, blue circles and blue diamonds). This point shall be returned to in the modelling (Section 7.5.3).

7.4.2 Peripheral tiger tail summation

Results from the two observers for the peripheral condition are shown in Figure 7.7. Within this region contrast sensitivity is relatively constant (Robson & Graham, 1981), and so the witch hat compensation has little effect on the shape of the summation slope. In general both observers show fourth-root summation as the stimulus size increases up to its maximum width of 32 cycles, which is consistent with the result reported by Robson and Graham (1981) for this location (Figure 7.3). Possible explanations for the steeper summation slope for the 1 cycle high stimuli in observer ASB will be explored in the modelling (Section 7.5.3).

The results presented in Figures 7.6 and 7.7 therefore show that for the foveal, parafoveal, and peripheral locations tested here (stimuli centred at 0, 12, and 42 cycles superior to the fixation point, equivalent to 0, 3, and 10.5 degrees of visual angle), fourth-root summation extended over 32 grating cycles (equivalent to 8 degrees of visual angle).

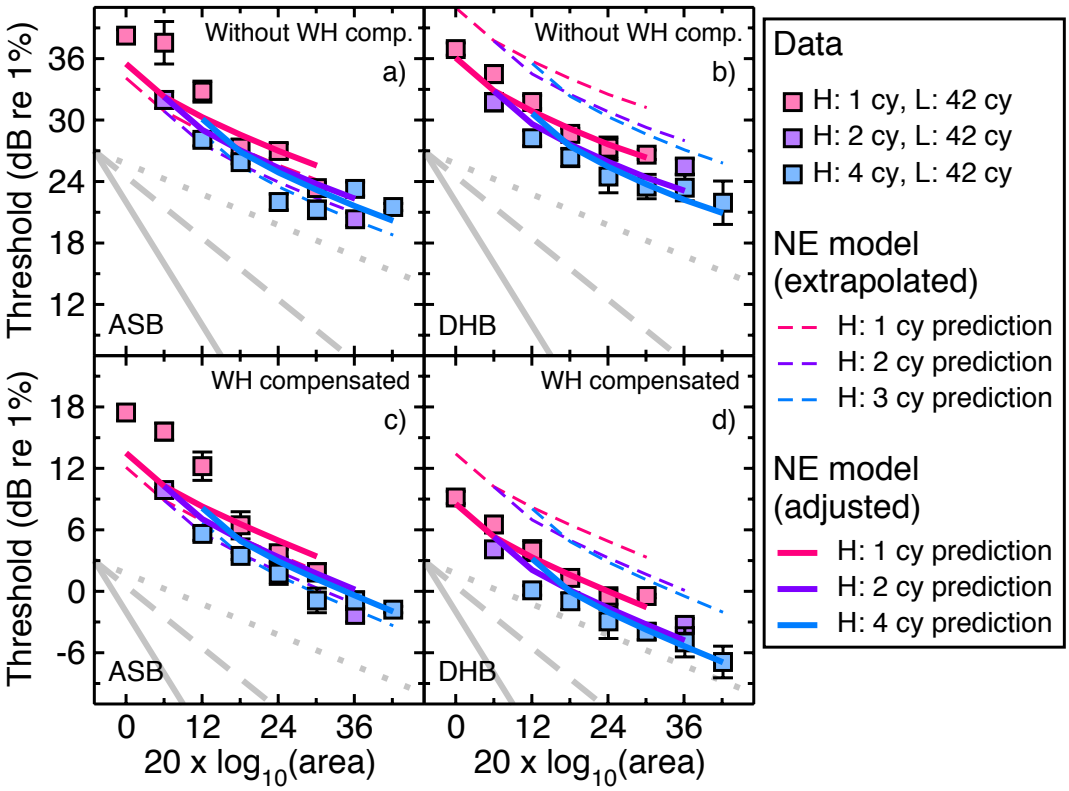

Figure 7.7: Contrast detection thresholds for "tiger-tail" stimuli presented in the periphery (centered at 42 cycles). The top two panels (a-b) show thresholds for the condition with the non-compensated stimuli, the bottom two panels (c-d) show thresholds for the witch hat compensated stimuli. The data are plotted with predictions made from the noisy energy model (NE) fit to the foveal and parafoveal data (dashed curves). Solid curves show direct model fits to the peripheral data (see Section 7.5).

7.5 Modelling

7.5.1 Model fitting

Two of the models tested in Chapter 6 are fitted to the data here. These are the noisy energy model (NE) and the noisy energy model with the flat template (FTNE). The difference between the two is that in the NE model the observers use a template matched to the expected signal exactly (including the attenuation with eccentricity), whereas in the FTNE model the template gives an equal weighting to the entire stimulus (a flat template). These models were chosen because the versions of these models with the template based on the stimulus extent (ENE and FENE) were the best fitting models tested in the previous chapter (see Section 6.5), and in this study the predictions from the NE and FTNE models are identical to those that would be made by the ENE and FENE models (the templates in each model are the same because unlike the "Swiss cheese" stimuli tested in the previous chapter, the tiger-tail stimuli do not have "holes" in them).

The NE and FTNE models are compared here in order to demonstrate the effect of having a matched template versus a flat template, and to evaluate the prediction made by each model against the data from this study. Although the distinction between the two models is subtle, they carry quite different theoretical implications. The matched template (NE) model involves an "ideal" combination of signal and noise over area by an observer that is aware both of the expected stimulus and of how the visual field inhomogeneity in contrast sensitivity will affect that stimulus. The flat template (FTNE) model, on the other hand, can represent an observer who simply ignores input from regions of the display that will not contain signal. Formally, both models implemented here assume the observer uses a template that follows the sinusoidal luminance modulation of the stimulus (rather than simply matching a template to the stimulus envelope), however this feature makes very little difference to the summation prediction as it is the effect of adding cycles to the stimulus that is under investigation here rather than the detail of what happens within a single cycle (Meese & Summers, 2012).

In the models tested here, the stimulus image is first multiplied by the witch hat attenuation surface as measured for each observer (see Section 5.2.1), and filtered by sine-phase log-Gabor filter elements (see Section 5.2.2). Each pixel of the output of this filtering is taken to represent the activity of a detector at that location. The behaviour of both models is then described by summing over the detectors

$$c_t = \sqrt{\frac{\sqrt{\sum_{i=1}^{n} (\sigma^2 \cdot t_i^2)}}{\sum_{i=1}^{n} (s_i^2 \cdot t_i)}}, \tag{7.1}$$

where c_t is the contrast detection threshold, s_i is the pixel value at each location in the attenuated and filtered image, σ_i is the standard deviation of the noise, and t_i is the template at that location. This is derived in Section 5.3.5. In the application of the model here it is assumed that the noise is constant ($\sigma_i = 1$) and Gaussian.

For a template which is matched to the transduced stimulus exactly ($t_i = s_i^2$), Equation 7.1 is the noisy energy model (NE). The flat template noisy energy model (FTNE) has templates (t_i) unaffected by the visual field inhomogeneity or witch hat compensation (the template always has a flat contrast profile). Summation slope predictions from the Minkowski summation approximation to the signal detection theory (SDT) probability summation model would be identical to those from the noisy energy model in this case (see Section 5.5.1). Predictions from the high threshold theory (HTT) probability summation model based on the psychometric slope would be similar, as the Weibull β for each observer was approximately 4. Regardless, as stated previously the HTT probability summation model has been thoroughly discredited (see Section 2.4.2; Swets et al., 1961; Green & Swets, 1966e; Nachmias, 1981; Laming, 2013).

The two models were fitted to the thresholds from the three observers, using the downhill simplex method (see Section 3.7) to minimise the RMS error between the data and the model

		Fovea & parafovea		Periphery	
Model	Observer	Offset (dB)	RMSe (dB)	Offset (dB)	RMSe (dB)
Noisy energy (NE)	ASB DHB SAW	13.53 14.85 11.47	0.79 0.95 1.51	1.39 -4.89	2.37 1.21
Flat template noisy energy (FTNE)	ASB DHB SAW	13.48 14.76 11.38	0.78 0.91 1.46	1.44 -4.81	2.38 1.21

Table 7.2: Vertical offset parameters and RMS errors for the noisy energy (NE) model fitted to the data in Figures 7.6 and 7.7, and the flat template noisy energy (FTNE) model fitted to the data in Figures 7.8 and 7.9. The fitted offset parameter in the periphery is the additional offset needed on top of that derived from the fits to the foveal and parafoveal data.

predictions. The only free parameter was the vertical offset of the model prediction curves (equivalent to varying the global sensitivity). In the first instance only the data for the foveal and parafoveal conditions were fitted. A prediction was then made for the peripheral condition based on the offset parameter from this fit. As the peripheral stimuli were presented beyond the region where the witch's hat was measured however, this prediction is based on extrapolating the attenuation surface beyond the region where it is empirically supported. To correct for this, an additional fit was performed solely to the peripheral data. The variation in contrast sensitivity over the stimulus display region at that eccentricity would be very minor (it was chosen for this reason by Robson & Graham, 1981), so an inaccuracy in the extrapolated witch's hat at this eccentricity would only cause a global over- or underestimation in sensitivity. This is entirely compensated for by refitting the offset parameter for this condition. The RMS errors and fitted vertical offset parameters for all of these fits are shown in Table 7.2.

7.5.2 Comparing the model predictions to the data

The noisy energy models with the matched template (NE) and the flat template (FTNE) are fitted to the foveal and parafoveal results from the three observers in Figures 7.6 and 7.8 respectively. The predicted and direct fits to the peripheral data are shown in Figures 7.7 and 7.9. The predictions from the two models are very similar, and this is reflected in the marginal differences in the RMS error that result from fitting them (see Table 7.2). The major difference between the two models is that the FTNE model predicts that thresholds should increase for the largest non-compensated stimuli presented in the fovea, as extending the template into these less sensitive regions results in the accretion of more noise than signal. The NE model on the other hand predicts monotonic summation slopes, as the matched template downweights the contribution of the signal and noise from the more eccentric regions. Due to the similarity of the model predictions for these two architectures, it is not possible to distinguish between them in this study. From this point, references will be made to the figures with the FTNE model

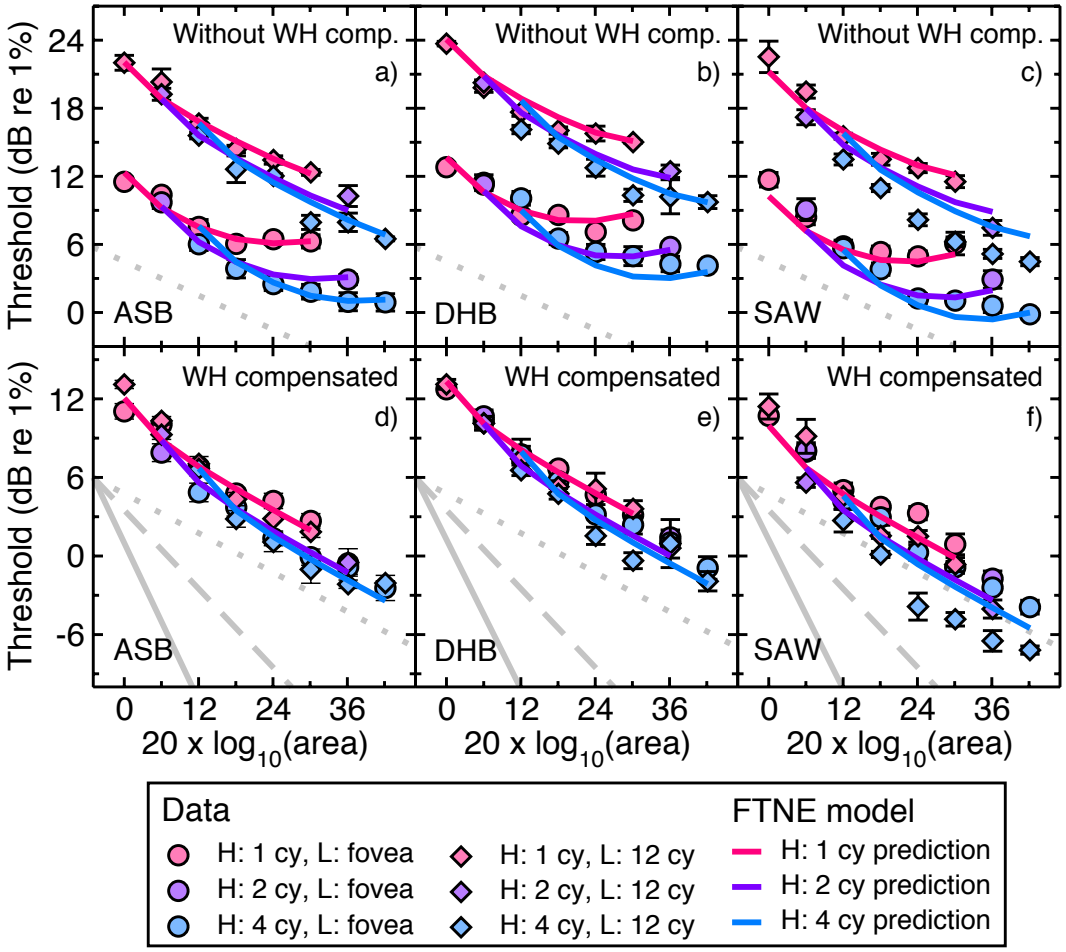

Figure 7.8: As Figure 7.6, but fitted with the flat template noisy energy model (FTNE).

fits (Figures 7.8 and 7.9), however in all cases the NE model fits were similar.

For all locations and observers, the noisy energy models provide a good fit to the data. The models predict the initial steep ($\approx -\frac{1}{2}$) decline in threshold from within-filter summation, the subsequent fourth-root summation slope, and the differences in sensitivity between the fovea and the parafovea. The models also capture the shallowing of the summation slope that results from the visual field inhomogeneity for the non-compensated stimuli (Figure 7.8a-c). For the witch hat compensated stimuli (Figure 7.8d-f) the models predict that results from the different locations should be identical, as the effects of the visual field inhomogeneity have been factored out. The empirical results are in agreement with this prediction (with the exception of one condition for observer SAW, discussed below).

The prediction for the peripheral stimuli made using the extrapolated witch's hat for observer ASB is very close to that found by fitting the model directly to the peripheral data (cf. dashed and solid curves in Figure 7.9a and c). For observer DHB, the extrapolated witch's hat underestimated the sensitivity in the periphery (cf. dashed and solid curves in Figure 7.9b and d). This is consistent with the results presented in Section 4.8.1, which show that the true sensitivity

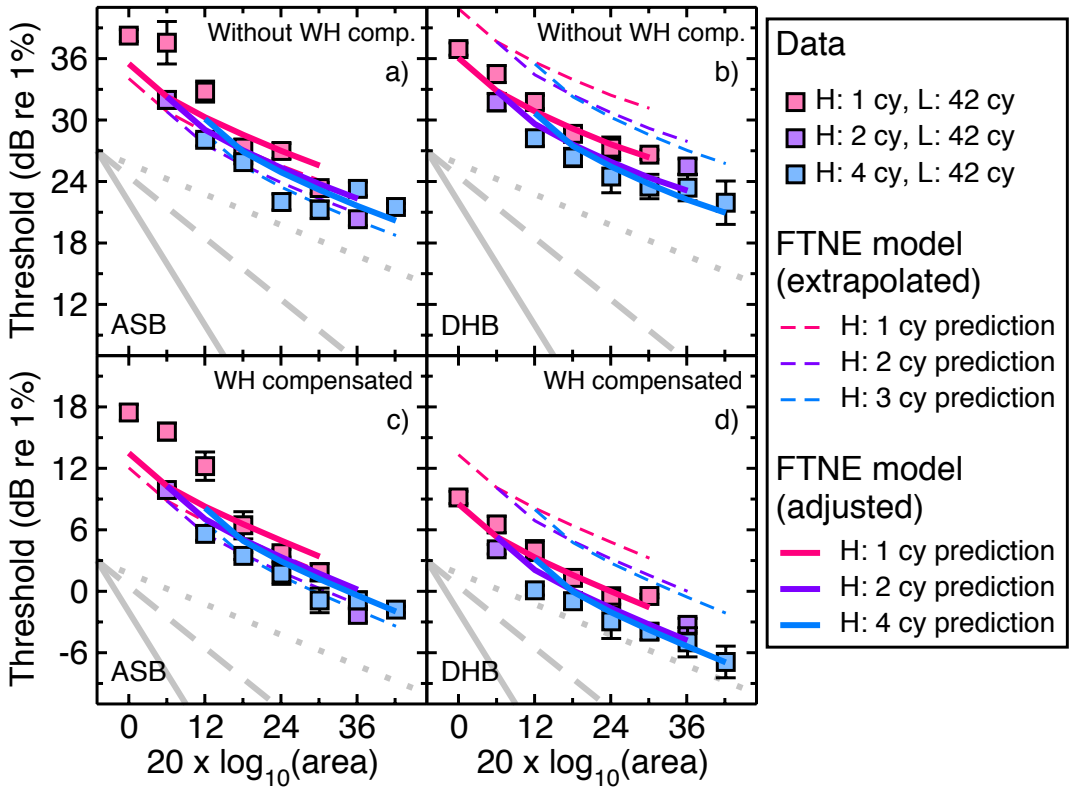

Figure 7.9: As Figure 7.7, but fitted with the flat template noisy energy model (FTNE).

decline beyond 18 cycles is slightly shallower than that predicted by the witch's hat.

7.5.3 Individual differences

In general, the model predictions provide excellent fits to the data. There are two cases however where observers show more summation than is predicted by either model. The first is found in observer SAW when 4 cycle high stimuli were presented in the parafovea. Thresholds for this condition were systematically lower than the model prediction. One possible explanation for this discrepancy would be if the receptive fields for this observer became elongated in the parafovea. This would result in more extensive linear within-filter summation. Another consequence of receptive field size covarying with eccentricity would be that the stimuli used to map the witch's hat attenuation surface (which were log-Gabors with constant bandwidths equal to those used in the filtering here) would overestimate the sensitivity decline. This is in agreement with the effect seen in Figure 7.8, where data from the 4 cycle high stimuli in the parafovea lie consistently *below* the model prediction curve.

Figure 7.10 shows a modified FTNE model fit to the data for observer SAW. The orientation bandwidth of the log-Gabor used for filtering the stimuli in the parafovea was reduced from $\pm 25°$ to $\pm 14.7°$, elongating the receptive fields along the grating stripes. This parameter was adjusted by hand to fit the data. As the assumption of equal receptive field size underlying the

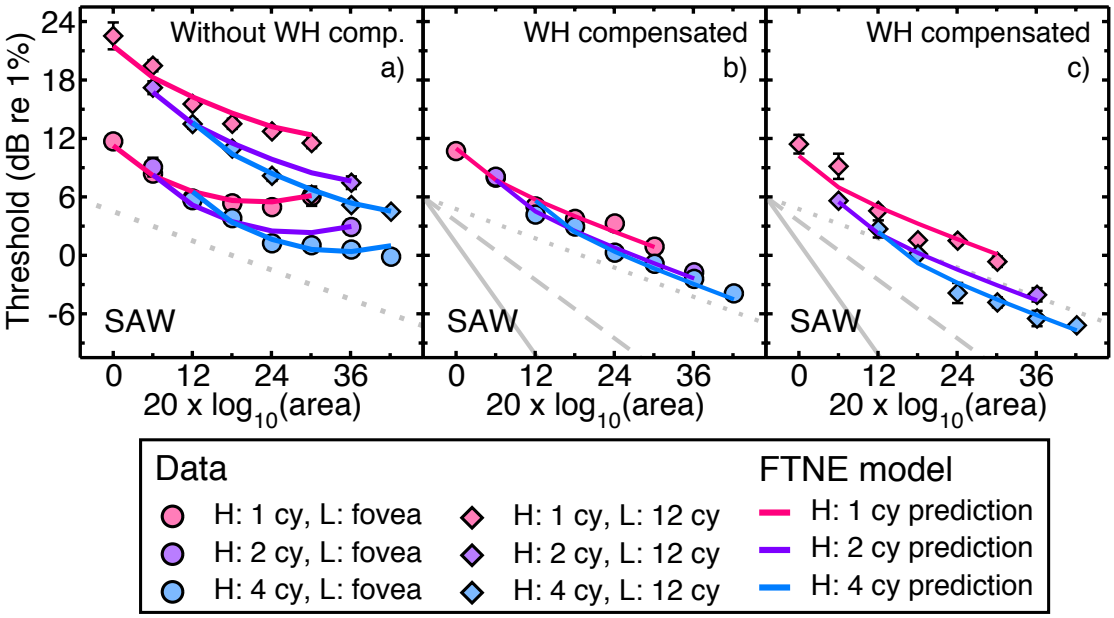

Figure 7.10: A modified version of the FTNE model fitted to the foveal and parafoveal data for observer SAW (cf. Figure 7.8). For clarity, the data from the two locations are shown in separate subplots (b-c) for the witch hat compensated stimuli. The orientation bandwidth of the log-Gabor used in the filtering stage for the parafoveal stimuli was narrower ($\pm 14.7°$), elongating the simulated receptive field. The foveal and parafoveal data were fitted with separate vertical offset parameters (10.25 dB for the foveal fit, 6.58 dB for the parafoveal fit). This model provides a superior fit to the data (RMSe = 0.71 dB).

attenuation surface was violated, the model curves were fitted with separate vertical offset parameters for the foveal and parafoveal data. The fit from this model (RMSe = 0.71 dB) was superior to that from the model with equal filter bandwidths (RMSe = 1.46 dB). This was also the case for the Akaike's Information Criterion (AIC, see Section 4.7.3) scores, which take account of the additional two parameters for the modified model when assessing the fit quality (lower scores are better, for the two models these were -13.18 vs. 23.19). A nested model hypothesis test found the improvement in the quality of the fit provided by the addition of these two parameters to be highly significant ($F_{2,53} = 85.56, P < 0.001$).

For observer ASB, thresholds for the smallest stimuli (1 cycle high, 1-4 cycles wide) were approximately 3 dB higher in the periphery than the models predicted (Figure 7.9a and c). Receptive field elongation would not explain this result, as the sensitivity offset this predicts for different stimulus heights would apply for *all* stimulus widths. One possible explanation would be if the observer could not use a template matched to the stimulus for these conditions. This would make the amount of noise integrated in the summation process constant across the smaller sizes, therefore steepening the summation slopes to that predicted by a quadratic summation model. This causes the data to lie *above* the model prediction because the witch's hat was not sampled in this region and therefore the effect is not built into the attenuation surface. Such a limitation on the template stage may result from uncertainty about

the location of the stimulus in the periphery, or mandatory signal combination over a minimum summation region (i.e. crowding; see Parkes et al., 2001).

7.5.4 The stochastic max-across-templates model

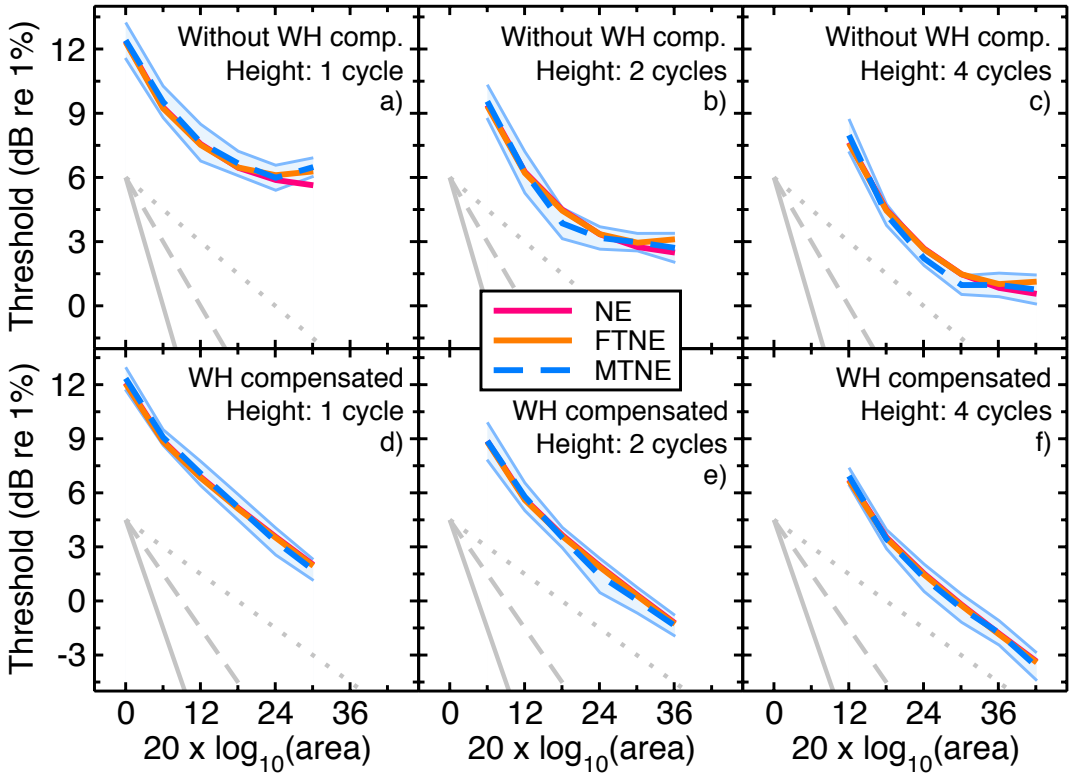

Figure 7.11: Stochastic Monte-Carlo model predictions generated for the max-across-templates noisy energy (MTNE) model. Average thresholds and the standard deviation of the threshold estimate (blue shaded region) were calculated from 10 repetitions of a simulated experiment (50 trials per point on every psychometric function in each repetition), using the witch hat attenuation surface from observer ASB. The MTNE model is plotted with the NE and FTNE models.

The complication for applying matched template models to experiments which interleaved different stimuli (discussed in Section 6.5.4) is revisited here. As in the previous chapter, the non-compensated and witch hat compensated stimuli were interleaved in this study. This complicates the prediction from the NE model for this condition, as the observer would not know on each trial whether to use the template matched to the non-compensated or to the witch-hat compensated stimuli. As a result, the analytic formulation of the NE model would slightly overestimate performance. The FTNE model does not have such a problem, as the templates used in this model were identical for the two conditions.

A more plausible version of the NE model would be one that monitored the templates for both types of stimulus in each interval, and then used the template most likely to be matched to the

stimulus. This is the method proposed by Meese and Summers (2012), who used a max operator to select the more activated template (after normalising by the standard deviation of the expected combined noise in each template). The analytic derivation of such a model would be quite difficult however, in this case it is simpler to generate a stochastic prediction. Figure 7.11 shows predictions for this max-across-templates noisy energy (MTNE) model compared against the predictions from the NE and FTNE models. The predictions from the MTNE model are very similar to those from the NE and FTNE model. It is apparent that the stimulus conditions tested in this study could not clearly distinguish between the three models.

7.6 Discussion

7.6.1 Summation behaviour is the same across the visual field

Once the effects of the visual field inhomogeneity are compensated for, summation behaviour becomes equivalent across the visual field. For all observers the results were well-described by the noisy energy model. It is recognised however that the results of this study could also be fit by a probability summation model, which is rejected for other reasons (see Meese & Summers, 2012). Whether the template stage in the model features matched or flat templates, and the way in which these templates were selected, made little effect on the prediction for these results. Distinguishing between these model variants would require a study specifically designed to exaggerate the differences in the predictions.

The visual field inhomogeneity in contrast sensitivity (see Chapter 4) will have prevented previous studies from being able to find this result when presenting stimuli to the fovea. Although some data from previous foveal studies have a fourth-root summation slope (e.g. some observers and conditions in Robson & Graham, 1981; Polat & Norcia, 1998; Polat & Tyler, 1999; Meese & Hess, 2007) in most cases summation slopes become shallower as the stimulus size increases (see Figure 7.1). Without the compensation used here for the visual field inhomogeneity it has not previously been possible to demonstrate that summation behaviour in the fovea, the parafovea, and the periphery follows the same rule.

The fourth-root summation found in the periphery here agrees with the results of some previous studies (see Figure 7.3; Robson & Graham, 1981; Meese & Hess, 2007). The within-filter summation between the smallest stimulus sizes (1 and 2 cycles wide) steepens the summation slope to the $-\frac{1}{2}$ gradient found in some other studies that tested in the periphery, but not for stimuli of comparable size (Figure 7.3; Manahilov et al., 2001; Meese & Hess, 2007). An alternative explanation for these results is given below. I found no evidence to support the linear summation found in the periphery by Mayer and Tyler (1986). It is not clear whether this aberrant result can be reconciled with the findings from other studies.

7.6.2 Receptive field elongation and template effects

The results from two of the observers do show small differences in behaviour under the different location conditions. In both cases, more summation than predicted by the basic version of the models is shown at the more eccentric locations. For observer SAW this increase in summation is explained by an elongation of the receptive fields in the parafovea. Building this effect into the model does then equalise summation behaviour between the different locations. Such a modification could be justified by reference to animal physiology, where increases in receptive field size with eccentricity have previously been found (Wilson & Sherman, 1976). The two other observers do not show any evidence of receptive field elongation with eccentricity.

For observer ASB, the increased summation found in the periphery for the smallest stimuli cannot be explained by changes in receptive field size. One possible explanation for this result would be a limitation on template stage such that matched templates could not be formed for these stimuli in the periphery. If this were the case the internal noise would remain constant across the smallest stimulus conditions, increasing the rate of threshold decline (effectively turning the noisy energy model into the quadratic model for those stimuli, see Section 5.3.3). Problems with the template stage could also explain the quadratic summation found in some previous studies that tested small stimuli in the periphery (Manahilov et al., 2001; Meese & Hess, 2007). Although Manahilov et al. (2001) attempted to reduce uncertainty by presenting a suprathreshold stimulus example before each block, is is possible that the observer would not be able to maintain a template at the correct size and location from memory alone. Further work is required to determine whether a breakdown in the template stage can explain the increased summation found under these conditions.

CHAPTER 8

Battenberg summation in amblyopes

8.1 Motivation and summary

Previous chapters (6 - 7) have investigated area summation in the normal visual system, where I have been able to predict behaviour under a range of stimulus conditions. One case in which the visual system behaves abnormally is the clinical condition of amblyopia. Previous experiments conducted on amblyopes have found higher thresholds for contrast detection, but normal summation behaviour. At the time this was accounted for by the same probability summation model as was popular for normal vision. Recent work has suggested that a better explanation of area summation in normal vision is provided by a noisy energy model. This raises the question of whether the interpretation of summation in amblyopes should also be updated.

The "Battenberg" study conducted by Meese (2010) was designed to differentiate between the predictions from probability summation and noisy energy models. Stimuli were designed such that summation could be measured without the confounding effects of the visual field inhomogeneity in contrast sensitivity and whilst keeping the overall stimulus extent constant. This study found more summation in normal observers than would be possible under probability summation, instead agreeing with a noisy energy model. I have repeated this experiment with amblyopic observers to investigate summation behaviour in amblyopia, and determine which model provides a better account. My findings suggest that although thresholds are higher in the amblyopic eyes of amblyopes, summation behaviour is normal in the two eyes.

8.2 Introduction

8.2.1 Battenberg summation in the normal visual system

The study of area summation is complicated both by the visual field inhomogeneity in contrast sensitivity (as described in Chapter 4), and the accumulation of additional noise as the stimu-

lus size increases (if the observer can adjust the monitored region of the visual field). Meese (2010) introduced Battenberg stimuli to counteract both of these effects (described in Section 3.4.6 and shown in Figure 8.1). Briefly, these stimuli are formed from arrays of micropattern elements. Each element is a single cycle of a sine wave carrier grating windowed by an orthogonal cosine half-cycle at half the spatial frequency of the carrier. Varying the number and position of the elements in the display allows for area summation behaviour to be measured across sets of stimuli with a constant extent. This mitigates the confounding effects of both the inhomogeneity and the variation in the amount of integrated noise for stimuli of different sizes, allowing the measured threshold difference to reflect only the effect of adding additional signal to the stimulus.

In the Battenberg stimuli the micropattern elements are arranged in a "checkerboard" pattern of regions containing contrast and regions which are blank. Each checked Battenberg stimulus contains approximately half of the total contrast over its area as the full stimulus. Comparing thresholds for the checked Battenberg stimuli to those for the "full" stimulus therefore allows the level of area summation to be measured and expressed as a summation ratio (this can be expressed either directly as a ratio, or in dB). By adjusting the scale of the checkerboard pattern the Battenberg stimuli can be used to measure local summation occurring over short distances (smaller check sizes), or long-range summation occurring over greater distances (larger check sizes).

The level of summation found for smaller check sizes by Meese (2010) was 6 dB (thresholds doubled for the checked Battenberg stimuli), which can be explained by short-range linear summation within the receptive fields of simple cells (sometimes also referred to as "within-filter" summation). For the larger check sizes, 3 dB of summation was found ($\sqrt{2}$ increase in threshold between the full and Battenberg stimuli). This long-range summation is greater than would be possible under probability summation (see Section 5.4.3). It would be consistent however with the quadratic summation model (Section 5.3.3), the ideal summation model (Section 5.3.4), or the noisy energy model (Section 5.3.5) in the case where the observer was unable to match a template to the checkerboard stimulus pattern and instead defaulted to summing over the entire stimulus. Previous work presented elsewhere (e.g. Meese, 2010; Meese & Summers, 2012) and in this thesis (Chapters 6 - 7) supports the noisy energy interpretation of the Meese (2010) result.

8.2.2 The neural deficit in amblyopia

Amblyopia is an acquired disorder of the visual system that results in a permanent poor quality of vision in an eye. It is caused by an impairment in that eye during the development of the visual system in childhood (usually caused by either strabismus or anisometropia). Previous research suggests that the amblyopic deficit is due to abnormal development of the

binocular cells in visual cortex (see Section 2.8.1; Hubel & Wiesel, 1970; Blakemore & Vital-Durand, 1986; Anderson et al., 1999; Barnes et al., 2001). Observers using their amblyopic eye to perform a task typically exhibit poor acuity, higher contrast detection thresholds (particularly for high spatial frequency stimuli), and perceptual distortions (see Section 2.8.2). The non-amblyopic ("fellow fixing") eye of an amblyope is thought to be functionally normal, and is often used as a convenient control (e.g. Hess et al., 1978).

The normal visual system is understood to perform spatial filtering operations on stimuli and represent the information from those stimuli in channels which are tuned to different spatial scales (see Section 2.3.2; Campbell & Robson, 1968). The existence of these channels is evidenced by adaptation studies (Blakemore & Campbell, 1969), and similar studies conducted in amblyope observers reveal that they are still present in amblyopia (Hess, 1980). The sensitivity for each spatial frequency can be determined by measuring a contrast sensitivity function (CSF). The manner in which amblyopia affects the CSF shows considerable variation across observers, prompting some authors to suggest that there may be separate types of amblyopia with different functional effects (Hess & Howell, 1977). Some amblyope observers show no loss of sensitivity, some a loss of contrast sensitivity only to higher spatial frequencies, and others a loss across the entire spatial frequency range (Hess & Howell, 1977; Levi & Harwerth, 1977; Hess et al., 1978; Thomas, 1978; Levi, Harwerth, & Manny, 1979).

8.2.3 Area summation in the amblyopic visual system

For the summation of luminance-defined stimuli, amblyopic vision has been found to behave like the peripheral vision of normal observers (Flynn, 1967). This is consistent with ideal summation in the smallest available receptive field size, if the both the normal periphery and the amblyopic fovea suffer from a loss of detectors tuned to higher spatial frequencies (see Section 2.6.2). With contrast-defined (luminance modulated) stimuli such as a sine wave grating, area summation has been demonstrated to be normal in amblyopes (Hess & Howell, 1978; Hess & Campbell, 1980). These summation results were explained at the time by the probability summation model that was typically used to account for summation in the normal visual system (see Section 2.5.5). Second-order (contrast modulation) summation has also been demonstrated to be similar to that seen in normal observers (Wong & Levi, 2005).

8.2.4 This study

Previous research has found area summation over contrast-defined stimuli to be normal in amblyopes, and explained by probability summation. As the probability summation account of the behaviour of normal observers has recently been challenged however (Meese, 2010; Meese & Summers, 2012), amblyope summation behaviour was re-examined here using the

Battenberg stimulus paradigm. This stimulus design has previously been successful in distinguishing between probability summation and noisy energy model predictions, in the case where the observer is not able to match a template to the checkerboard pattern in the Battenberg stimuli. The results indicate that area summation is normal in amblyopia. For three out of six observers tested, behaviour was best described by a noisy energy model. For the other three observers the results were inconclusive.

8.3 Methods

8.3.1 Equipment

Stimuli were stored in a CRS ViSaGe and presented on a gamma-corrected CRT monitor (Mitsubishi Diamond Pro 2070). The monitor had a refresh rate of 120 Hz, and a mean luminance of 52.4 cd/m^2. The monitor was viewed from a distance of 1.19 metres for the 4 c/deg stimuli (at this distance there were 48 pixels per degree of visual angle). This viewing distance was adjusted in order to present the stimuli at the appropriate spatial frequency for each observer (2.5 to 8 c/deg, corresponding to viewing distances of 0.74 to 2.38 metres).

8.3.2 Stimuli

This experiment used horizontal Battenberg stimuli composed of similar micro-pattern elements as those used to generate the "tiger-tail" stimuli in Chapter 7. The Battenberg stimulus design is described in Section 3.4.6, and the stimuli used in this study are shown in Figure 8.1. Stimuli were 42 by 42 cycle arrays of Battenberg elements. In addition to the full stimulus (which has a nominal check size of 0), five check sizes were used (having 1, 2, 4, 6 and 8 element wide checks). Stimuli were presented in both the "white" (signal in centre) and "black" (no signal in centre) check phases.

8.3.3 Observers

Data were collected from one normal (ASB) and six amblyope (S1 - S6) observers. Details for each amblyope observer are provided in Appendix C. ASB was tested at 2.5, 4, and 8 c/deg. The tested spatial frequency for each amblyope observer was decided based on a pilot experiment. Detection thresholds were obtained for the full (0) stimulus in each eye, first at 2.5 c/deg and then increasing in spatial frequency (from 4 to 6 c/deg) until a difference in threshold of at least 6 dB was found between the two eyes. Data are presented from all of the amblyope observers who were recruited for this experiment. The intention was to repeat each condition four times

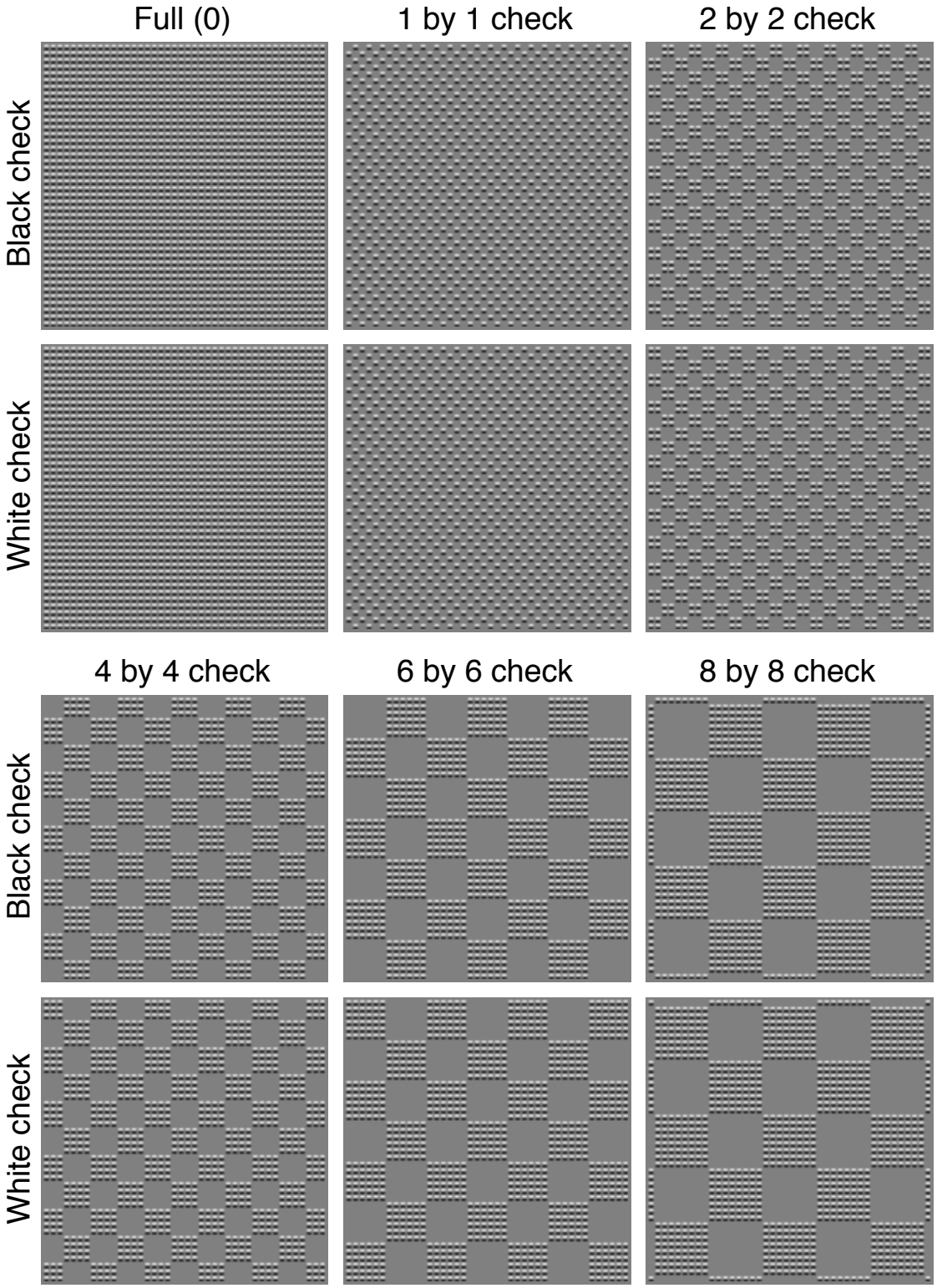

Figure 8.1: The "Battenberg" stimuli used in the experiments in this chapter. The full stimulus is shown, as well as the five different check sizes used in the experiments (check widths of 1, 2, 4, 6 and 8 elements). Each stimulus is shown in both the "black" and the "white" check phases.

for each observer, however due to time constraints this aim was not met and the number of repetitions varied across observer and condition (see Table 8.1).

Observers wore optical correction appropriate to the viewing distance. For amblyopes whose prescribed correction did not provide them with their best possible visual acuity, additional correction was added based on tests conducted in the lab. The experiments were carried out monocularly, alternating between the two eyes on each block. The non-tested eye in each block was occluded by an eyepatch. Observers gave informed consent before participation.

			Black checks						White checks					
		Check size	0	1	2	4	6	8	0	1	2	4	6	8
Observer	SF (c/deg)	Eye	Number of repetitions											
ASB	2.5	RE	4	4	4	4	4	4	4	4	4	4	4	4
		LE	4	4	4	4	4	4	4	4	4	4	4	4
ASB	4	RE	4	4	4	4	4	4	4	4	4	4	4	4
ASB	8	RE	2	2	2	2	2	2	2	2	2	2	2	2
		LE	2	2	2	2	2	2	2	2	2	2	2	2
S1	2.5	FFE	2	2	2	2	2	2	2	2	2	2	1	2
		AMB	3	3	3	3	3	3	3	3	3	3	3	3
S2	2.5	FFE	3	3	4	3	3	3	3	3	4	2	3	3
		AMB	3	3	3	3	3	3	3	3	3	3	3	3
S3	4	FFE	3	2	3	2	3	2	1	2	3	2	3	2
		AMB	2	2	2	2	2	2	2	2	2	2	2	2
S4	6	FFE	1	2	2	2	2	2	3	1	2	2	2	2
		AMB	2	3	3	2	2	2	2	2	2	2	2	2
S5	4	FFE	2	4	4	4	4	4	4	4	4	4	4	4
		AMB	3	4	4	4	4	4	3	4	4	4	4	4
S6	2.5	FFE	1	0	1	0	1	0	1	0	1	0	1	0
		AMB	1	0	1	0	1	1	1	1	1	0	1	1

Table 8.1: The number of repetitions used to generate the averaged data plotted for each observer and condition in this chapter. For the normal observer ASB, numbers of repetitions are shown separately for the right (RE) and left (LE) eyes. For the amblyopes the eyes are labelled as the fellow fixing eye (FFE) and the amblyopic eye (AMB).

8.3.4 Procedures

Thresholds were measured using a two-interval forced-choice (2IFC) three-down, one-up staircase procedure (see General Methods chapter, Section 3.5). For the normal observer (ASB) stimuli were organised into two blocks. In one block the full, 2 by 2 check, and 6 by 6 check conditions were interleaved (both black and white phases). In the other block, the 1 by 1, 4 by 4, and 8 by 8 check conditions were interleaved (both phases). For the amblyope observers these stimuli were reorganised into four smaller blocks in order to make each testing session

shorter. These were the full and the 2 by 2 check stimuli, the full and the 6 by 6 check stimuli, the 1 by 1 and the 4 by 4 check stimuli, and the 1 by 1 and the 8 by 8 check stimuli.

Feedback on correctness of response was provided. The contrast detection thresholds for each repeat were calculated using a Weibull fit to the staircase data in Palamedes (see General Methods chapter, Section 3.6), allowing the mean and standard error to be calculated across repetitions. Owing to the limited amount of time available for testing the amblyopic subjects, the rejection criterion based on the standard error from a Probit fit to the data (see Section 3.6.3) was relaxed to a maximum error of 5 dB in this study.

8.4 Modelling

8.4.1 Model architectures

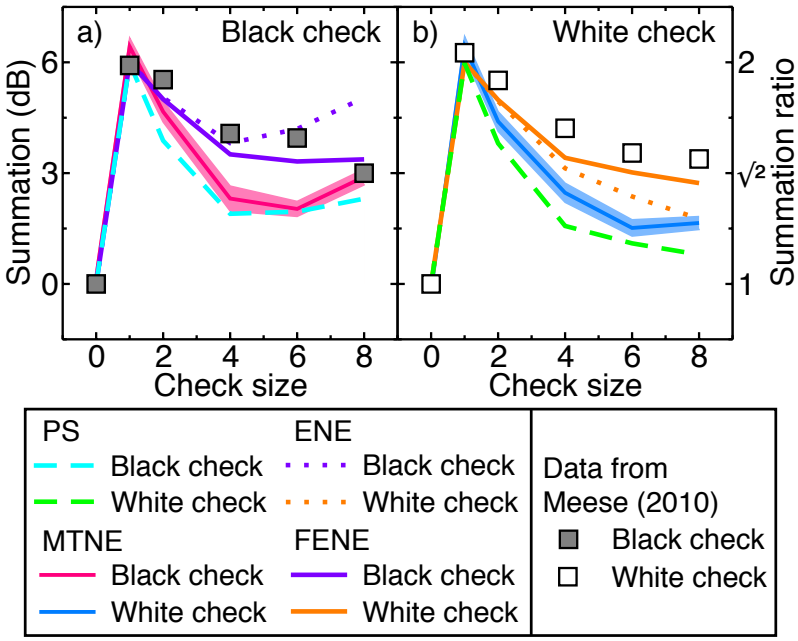

Figure 8.2: "Battenberg" summation data for normal observers reproduced from Meese (2010) and plotted with predictions from the probability summation model (PS), "max-across-templates" noisy energy model (MTNE), the "extent" noisy energy model (ENE), and the "flat extent" noisy energy model (FENE). The shaded area shows the standard deviation of the responses from the stochastic MTNE model. All model prediction curves, and the data from Meese (2010), were normalised so that the full (0) stimulus had a threshold of 0 dB. The spatial frequency used in Meese (2010) was 2.5 c/deg.

The models considered in this chapter are the signal detection theory (SDT) probability summation model (PS) and the "flat extent" noisy energy model (FENE). "Flat extent" refers to the nature of the template used to detect the stimulus, which is matched to the extent of the stimulus and has a constant weighting across its surface (it does not follow the checkerboard mod-

ulation in the stimuli, see Section 5.2.5). The PS and FENE models are respectively the previously proposed model for summation in amblyopia and the current favoured model for summation in the normal visual system. The aim of this study is to determine which model better characterises behaviour in amblyopia. The max-across-templates noisy energy model (MTNE) and the extent noisy energy model (ENE) are also presented in this section for completeness.

The FENE model is roughly equivalent to that found to provide the best account of the data in the previous Battenberg study (Meese, 2010), with the main difference being the inclusion of the bilinear witch's hat attenuation surface to account for the visual field inhomogeneity in contrast sensitivity (Chapter 4; Baldwin et al., 2012) as opposed to the linear decline in log sensitivity used in Meese (2010). Where only data from observer ASB were being fitted, the witch's hat parameters measured for that observer were used to generate the attenuation surface. Where data from other observers were fitted the attenuation surface used was that derived from fits to the averaged data from four observers (presented in Chapter 4).

There are two potential objections to the use of the witch's hat attenuation surface here. The first is that the surfaces are based on a binocular study, but their usage here is as a model of monocular sensitivity. The second is that these descriptions of the sensitivity of normal observers are used here to model amblyope behaviour. In response to the first objection, although it is possible that a monocular attenuation surface would feature asymmetries not present in the witch's hat (a nasal-temporal asymmetry would be predicted from the physiology, see Anderson et al., 1991), there is no reason to expect these asymmetries to affect the pattern of thresholds for the detection of symmetrical stimuli such as the ones used here. In response to the second objection, the stimuli are designed in such a way that the effects of the visual field inhomogeneity are largely mitigated (compare the predictions of the FENE model for the black and white check stimuli in Figure 8.2). For that reason it is unlikely that any distortions to the attenuation surface that might be caused by amblyopia would affect the findings.

Following the witch's hat attenuation of the stimulus image, all four models shown in Figure 8.2 include filtering by sine-phase log-Gabor filter elements (spatial frequency bandwidth of 1.6 octaves, orientation bandwidth of $\pm 25°$). The PS model then uses Minkowski summation to approximate probability summation over independently noisy local detectors (see Section 5.4.3). The MTNE model is based on a combination of square law transduction followed by a template matched exactly to the stimulus. As the different stimuli were interleaved in this study however, the observer could not know in advance which stimulus to use on each trial. For this reason, the stochastic implementation shown previously in Section 7.5.4 was used, in which the observer tests the squared filter responses against the templates for each of the possible stimuli and then selects the template which provides the best match by performing a max operation over the outputs from the templates.

The ENE and FENE models are the same as those shown in Section 6.5, and are equivalent to the NE and FTNE models when the stimuli do not have "holes" in them (i.e. for stimuli which

are not “Swiss cheeses” or Battenbergs). These feature square law transduction followed by a template stage where the template is matched to the stimulus *extent* (does not follow the contrast modulation present in the checked Battenberg stimuli). In the ENE model the template weighting decreases with eccentricity according to the decline from the attenuation surface (providing an ideal weighting at each location according to the expected signal to noise ratio), whereas in the FENE model the template is flat (constant weighting across the template).

8.4.2 Model predictions

In Figure 8.2, the predictions from each model are plotted as summation ratios. The summation ratio for each checked Battenberg stimulus is the difference in threshold between that stimulus and the full stimulus (which had approximately twice the signal area of the checked stimuli). This is equivalent to plotting predicted thresholds if the prediction were normalised to be at 0 dB for the full (0) stimulus. Although all four models predict the same initial 6 dB peak in summation with a doubling of signal area (from linear summation within the receptive fields), the predicted summation from each model diverges for the larger check sizes. The PS model predicts that summation should asymptote at 1.5 dB. The FENE model predicts that the level of summation should decrease to approximately 3 dB. Predicted summation levels are slightly higher for the black check stimuli (see Figure 8.2a) and slightly lower for the white check stimuli (see Figure 8.2a) due to the effects of the inhomogeneity in contrast sensitivity. The ENE prediction is similar to the FENE prediction whereas the MTNE prediction is more similar to the PS prediction. The inhomogeneity of the template in the MTNE and ENE models however causes their predictions for the two different check phases to diverge more than they do in the PS and FENE models. The data from Meese (2010) are best described by the FENE model, which is in agreement with the findings of that study.

8.4.3 Fitting the model predictions to the data

Model curves were fitted to the results from the normal and amblyope observers below using the downhill Simplex method to minimise the RMS error between the model prediction and the data (see Section 3.7). These fits were performed independently for each eye, using a single free parameter which controlled global sensitivity (equivalent to varying the standard deviation of the internal noise).

8.5 Results

8.5.1 Normal observer

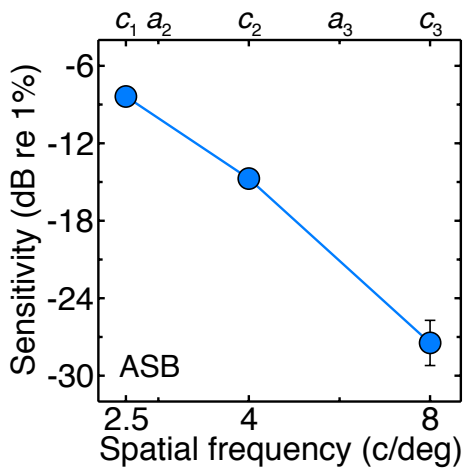

Figure 8.3: Detection thresholds for the "full" Battenberg stimuli for observer ASB at three spatial frequencies (c_1, c_2, and c_3). The labels a_2 and a_3 indicate the spatial frequency of the artefacts introduced by the 1 by 1 check Battenberg modulation in the 4 c/deg and 8 c/deg stimuli respectively (see Section 8.5.3).

A contrast sensitivity function (CSF) for observer ASB showing sensitivity at the three spatial frequencies at which this observer was tested (derived from the thresholds for the "full" Battenberg stimulus) is shown in Figure 8.3. Sensitivity is highest at 2.5 c/deg, and then declines for 4 c/deg and 8 c/deg. Figures 8.4 - 8.9 show Battenberg summation data from this observer at these three spatial frequencies. The results are plotted as summation, which is calculated by normalising the threshold data to the average threshold for the full (0) stimulus. Results are plotted for 2.5 and 8 c/deg for both the dominant (RE) and non-dominant (LE) eyes, as determined by the Miles test. For the 2.5 c/deg data (Figure 8.4) the results are very similar from the two eyes. For the 8 c/deg data (Figure 8.8) the results from the dominant eye (RE) are considerably more variable than those from the non-dominant eye (LE), as is evidenced by the much larger error bars. These data are from only two repetitions however (see Table 8.1), and the pattern of the mean thresholds is similar in the two eyes. At 4 c/deg data were collected only from the dominant eye (Figure 8.6).

For the 2.5 c/deg data (Figure 8.4) the results look similar to those found previously at this spatial frequency (Meese, 2010), as shown in Figure 8.2. Summation was around 6 dB (thresholds approximately doubled) for the smallest check size (1 by 1), before declining to approximately 3 dB of summation for the larger check sizes. Results at 4 c/deg were similar for the larger check sizes (Figure 8.6), however the initial peak in summation for the smaller check size stimuli is reduced to approximately 3 dB. The results from 8 c/deg are also similar for the larger check sizes, however at this spatial frequency the initial summation peak is even smaller.

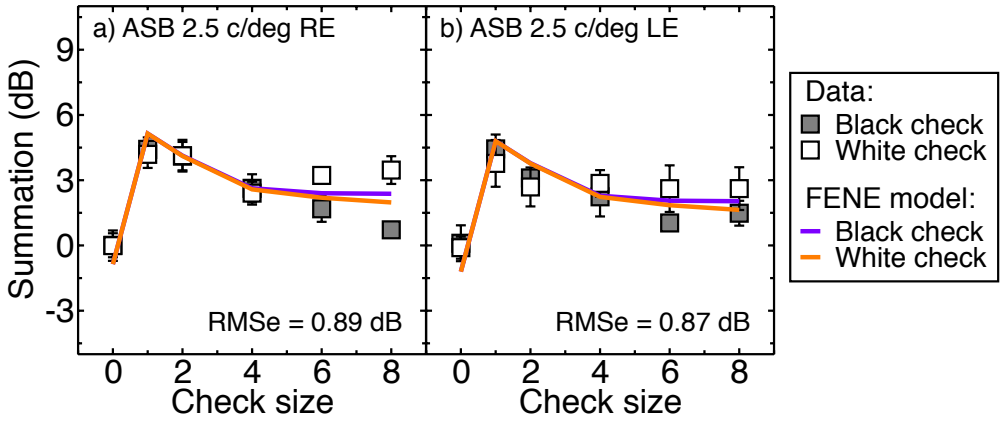

Figure 8.4: Detection thresholds from observer ASB for 2.5 c/deg "Battenberg" stimuli presented to the dominant (RE) and non-dominant (LE) eyes. The data are fitted by the prediction from the FENE model.

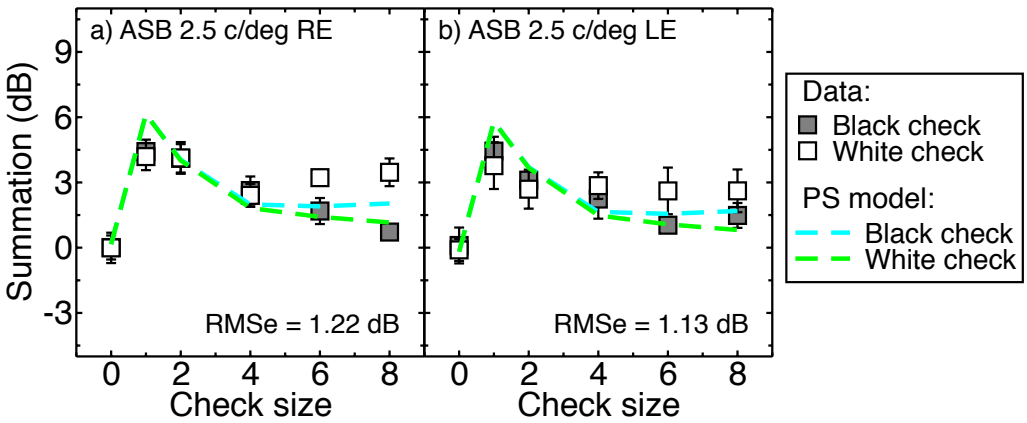

Figure 8.5: Detection thresholds replotted from Figure 8.4. The data are fitted by the prediction from the PS model.

			FENE model		PS model	
Observer	SF (c/deg)	Eye	Figure	RMSe (dB)	Figure	RMSe (dB)
ASB	2.5	RE	8.4	0.89	8.5	1.22
		LE		0.87		1.13
ASB	4	RE	8.6	1.31	8.7	1.64
ASB	8	RE	8.8	1.69	8.9	2.16
		LE		1.89		2.29

Table 8.2: RMS errors between the data from observer ASB and the FENE and PS model predictions for the fits shown in Figures 8.4 - 8.9. The rows labelled RE and LE refer to the data collected from the right and left eyes respectively.

Analysis presented later in this chapter indicates that the loss of the initial summation peak for higher spatial frequency stimuli might be due to artefacts in the Battenberg stimuli con-

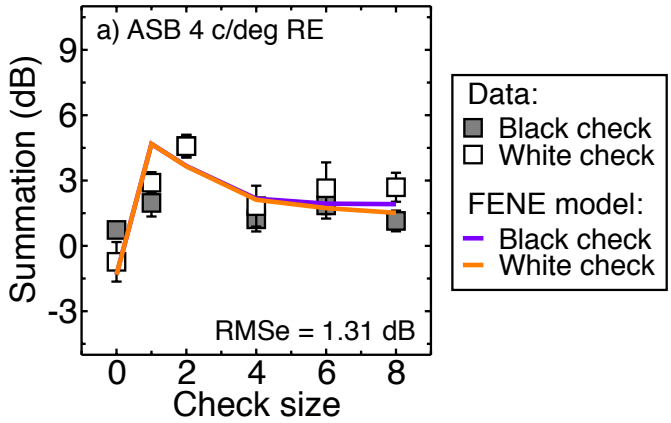

Figure 8.6: Detection thresholds from observer ASB for 4 c/deg "Battenberg" stimuli presented to the dominant (RE) eye. The data are fitted by the prediction from the FENE model.

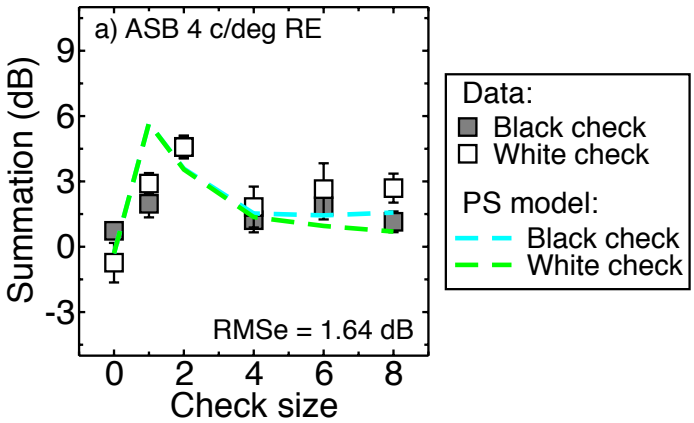

Figure 8.7: Detection thresholds replotted from Figure 8.6. The data are fitted by the prediction from the PS model.

tributing energy at spatial frequencies other than the carrier (this is discussed in Section 8.5.3). The observer could then use information from these artefacts to assist in the detection of the checked Battenberg stimuli, thus obscuring the magnitude of the full summation effect. The lack of such an effect in 2.5 c/deg data and in the previous Battenberg study (Meese, 2010) could be explained if stimuli were presented at the most detectable spatial frequency there.

The pairs of Figures 8.4 - 8.5, 8.6 - 8.7, and 8.8 - 8.9 each show the same data sets fitted by two different model predictions (those being from the FENE and PS models). The shapes of the model predictions differ slightly from those shown in Figure 8.2 because the witch's hat attenuation surface in the model for the predictions shown here is based only on data from observer ASB. In all cases the predictions from the FENE model provided a better fit to the data* than the PS model predictions (see Table 8.2). The linear filtering stage in the models causes them to both predict that for the smallest check size there should be a 6 dB increase in threshold from that for the full condition. This results from linear summation within the simple cell receptive fields being represented at the initial filtering stage in the models. As discussed

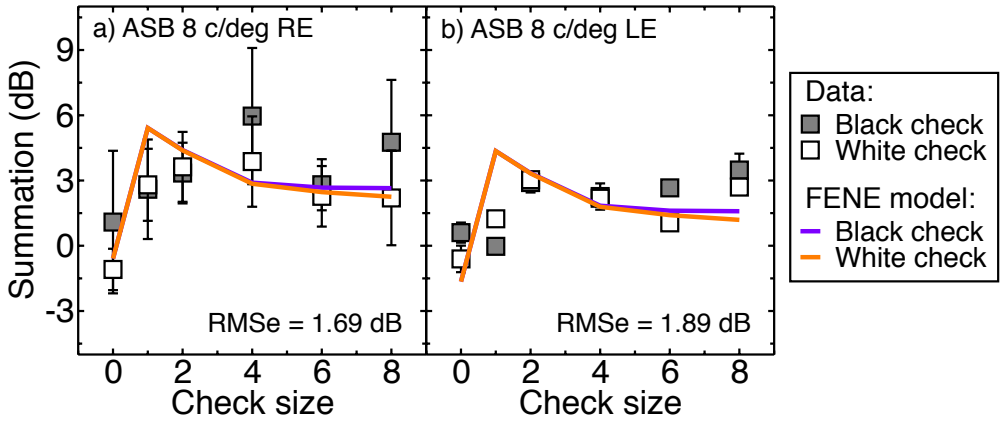

Figure 8.8: Detection thresholds from observer ASB for 8 c/deg "Battenberg" stimuli presented to the dominant (RE) and non-dominant (LE) eyes. The data are fitted by the prediction from the FENE model.

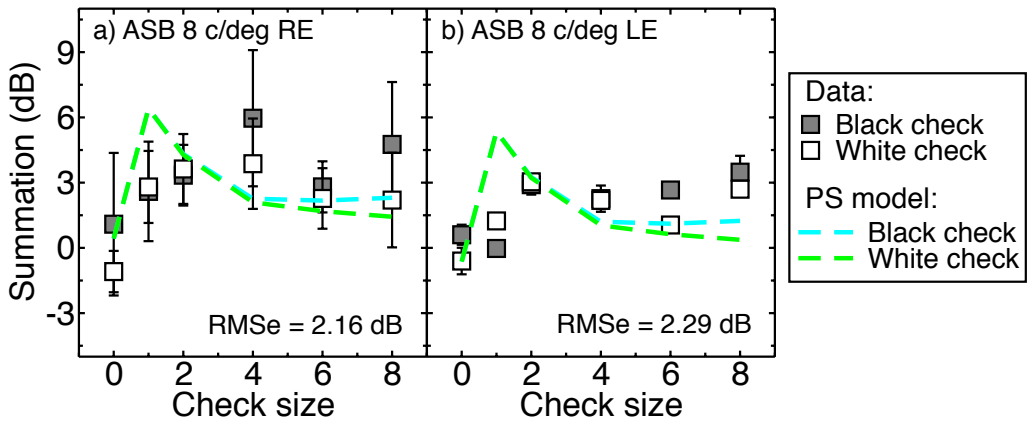

Figure 8.9: Detection thresholds replotted from Figure 8.8. The data are fitted by the prediction from the PS model.

above, the peaks in the data are accurate to those predicted by the models for the 2.5 c/deg stimuli but smaller than predicted for the 4 and 8 c/deg stimuli.

After the initial 6 dB peak for the smallest check size, which is explained by linear summation within receptive fields, summation decreases to approximately 3 dB. It is the summation level for these larger check size stimuli which is critical for preferring the FENE model over the PS model, as the FENE model predicts 3 dB of summation for the larger check sizes whereas the PS model predicts 1.5 dB of summation. The superiority of the fit for the FENE model in each case reflects that the level of long-range summation found is closer to that predicted by the FENE model than that by the PS model, which is to say that there is too much summation in the data for the probability summation account to prevail. For the larger check sizes the thresholds for the white and black check stimuli are similar in most cases. There is some separation in

*The fitting of the model predictions in this study (allowing sensitivity to vary) is more generous to the PS model than that used by Meese (2010), who normalised the prediction to the threshold for the "full" stimulus. In this analysis the fit of the PS prediction can be improved by the fitting procedure overestimating the threshold for the full stimulus (reducing the difference between the models regarding the prediction for the checked stimuli).

the right eye for the 2.5 c/deg stimuli (Figure 8.4a), where (contrary to the model predictions) thresholds are higher for the white check stimuli. The cause of this result is not clear.

8.5.2 Amblyopic observers

The threshold differences between the two eyes of the amblyopic observers tested in this study are shown in Figure 8.10a. For two observers (S4 and S5) these differences are less than the 6 dB sought when choosing the target spatial frequency. For these observers performance in the amblyopic eye was better relative to the normal eye on the main experiment than it was in the pilot experiment used to select the target spatial frequency. Figure 8.10b shows the differences in summation between the two eyes, calculated by subtracting the fellow fixing eye summation levels from the amblyopic eye summation levels. The data show that on average there is little difference in summation between the two eyes for the smallest (1 by 1) or larger (4 by 4 - 8 by 8) stimuli and 3 dB more summation for the fellow fixing eye for the 2 by 2 check stimuli. The summation results from the different amblyope subjects show considerable variability however, so in this section the data from each observer will be looked at individually and compared against the predictions from the FENE and PS models.

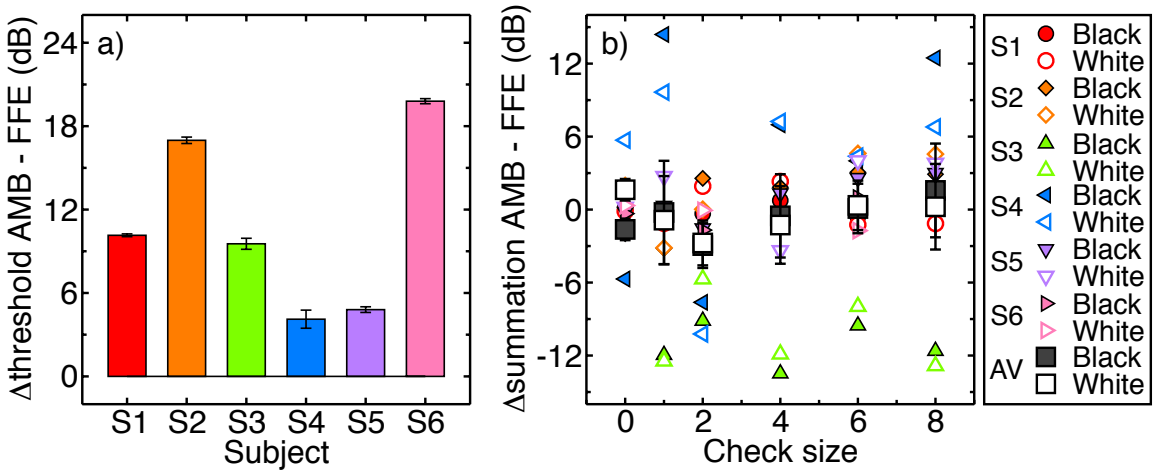

Figure 8.10: Panel a) shows the average threshold differences between the two eyes for each amblyope (S1 - S6) across the stimulus conditions tested here. Panel b) shows the difference in summation for each check condition, equivalent to the threshold difference between the two eyes normalised to that for the "full" (0) stimulus.

Figures 8.11 - 8.23 show summation results from the six amblyopic observers. Data are presented from the fellow fixing (FFE) and amblyopic (AMB) eyes in the two panels a) and b). The figures corresponding to each observer and model prediction are detailed in Table 8.3. For observer S4, thresholds for two conditions in the fellow fixing eye were much higher than those in the other conditions (see Figure 8.17a). These anomalous data were removed and the FENE model prediction was refit (Figure 8.18). This restricted data set was also fitted by the PS model (Figure 8.19). As above, the key distinction between the FENE and PS model

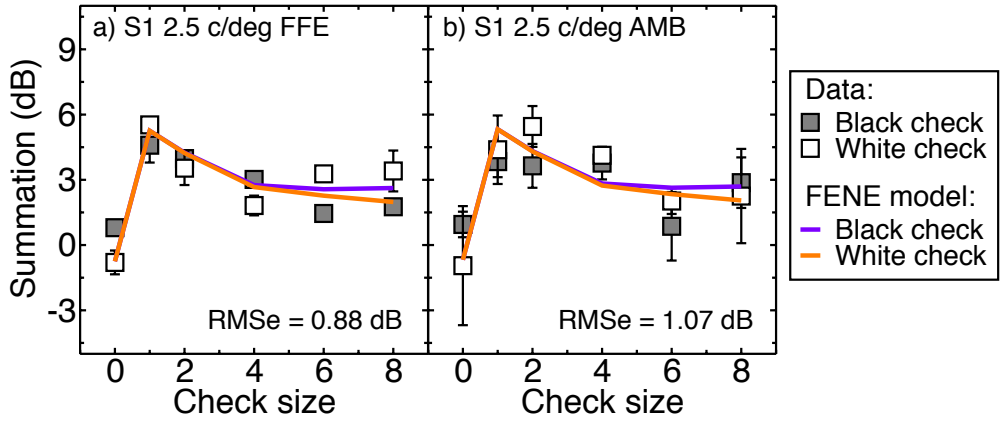

Figure 8.11: Detection thresholds from amblyopic observer S1 for 2.5 c/deg "Battenberg" stimuli presented to the fellow fixing (FFE) and amblyopic (AMB) eyes. The data are fitted by the prediction from the FENE model.

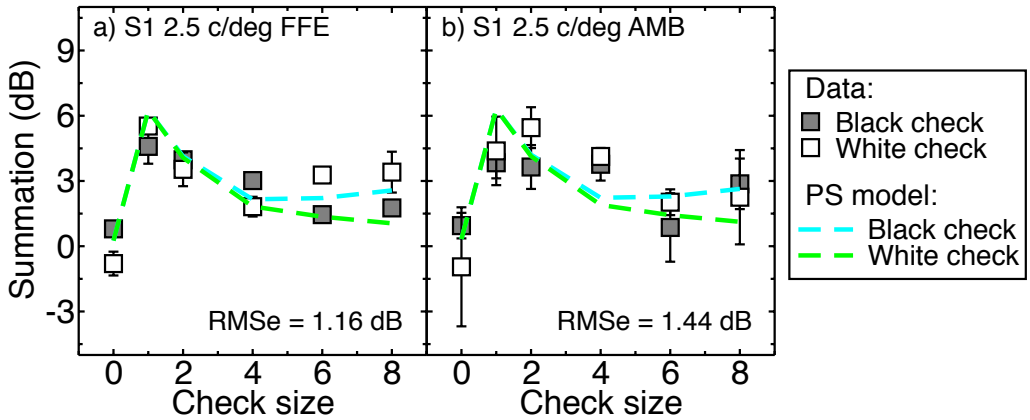

Figure 8.12: Detection thresholds replotted from Figure 8.11. The data are fitted by the prediction from the PS model.

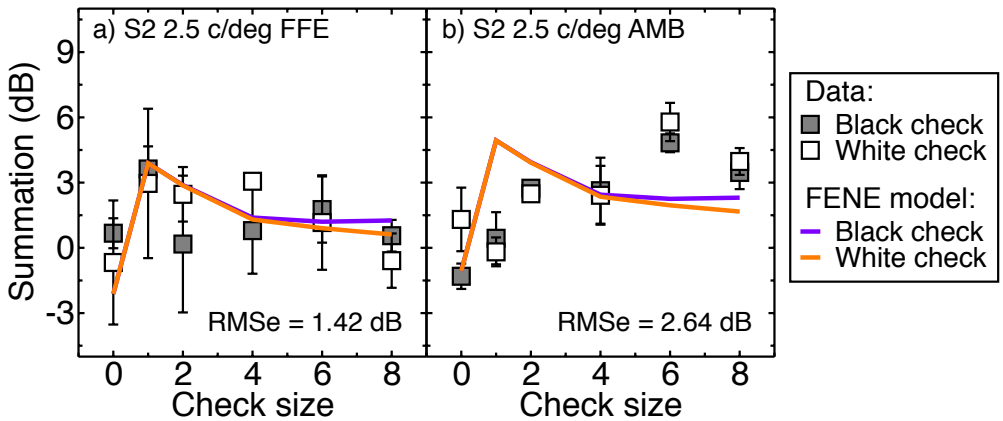

Figure 8.13: Detection thresholds from amblyopic observer S2 for 2.5 c/deg "Battenberg" stimuli presented to the fellow fixing (FFE) and amblyopic (AMB) eyes. The data are fitted by the prediction from the FENE model.

predictions here is that for the FENE model summation should asymptote at approximately 3 dB for the largest stimulus size, whereas for the PS model summation should asymptote at approximately 1.5 dB.

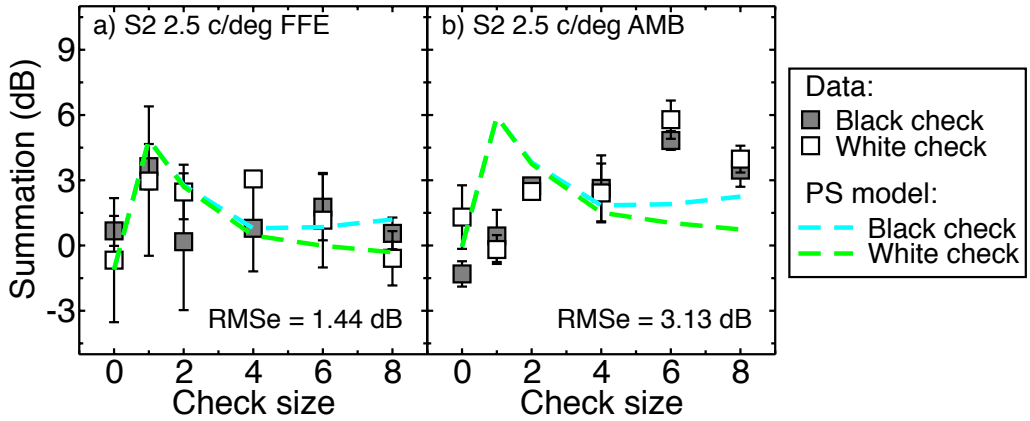

Figure 8.14: Detection thresholds replotted from Figure 8.13. The data are fitted by the prediction from the PS model.

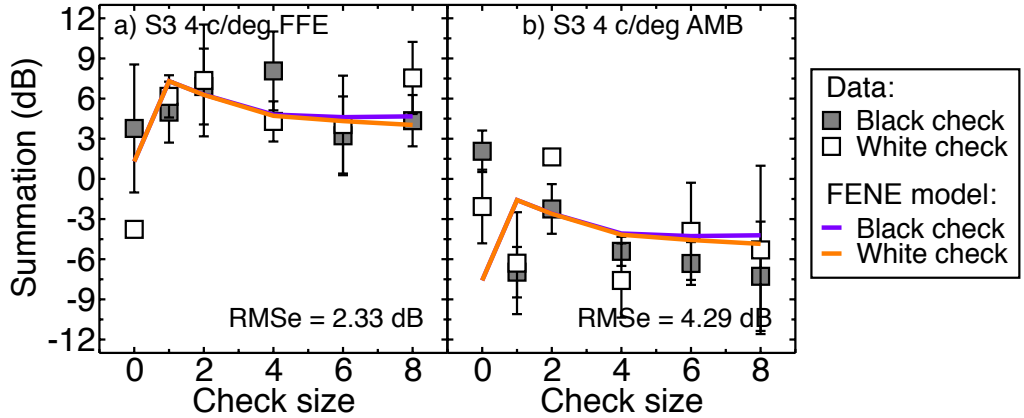

Figure 8.15: Detection thresholds from amblyopic observer S3 for 4 c/deg "Battenberg" stimuli presented to the fellow fixing (FFE) and amblyopic (AMB) eyes. The data are fitted by the prediction from the FENE model.

If the fellow fixing eyes of amblyopes are functionally normal, it would be expected that amblyopes using those eyes should behave in a similar manner to the normal observer tested in this study (ASB), and to those tested by Meese (2010). The fellow fixing eyes of observers S1, S5, and S6 provide data that are similar to previous results. Data from the other observers (S2, S3, and S4) do not clearly show the stereotypical pattern of threshold elevation for the Battenberg check conditions. For observer S3 (Figure 8.15) the initial summation peak is smaller than that predicted from linear summation within receptive fields. For the only amblyope tested at 6 c/deg (S4, see Figure 8.18) this initial summation peak is missing entirely. This may be due to the same problems with stimulus artefacts that I suggest caused a similar effect in the normal observer (see Section 8.5.3). In all cases however the data from the fellow fixing eyes of the amblyopes were fit better by the FENE prediction than by the PS prediction (see Table 8.3).

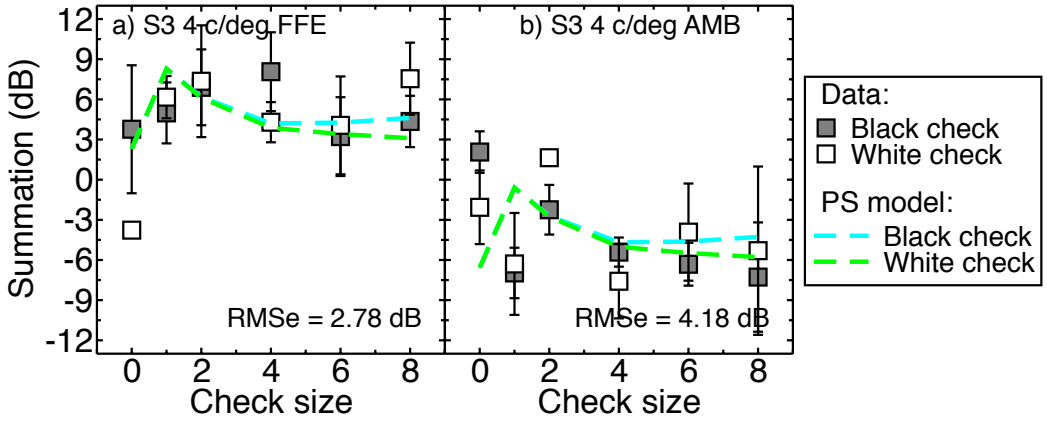

Figure 8.16: Detection thresholds replotted from Figure 8.15. The data are fitted by the prediction from the PS model.

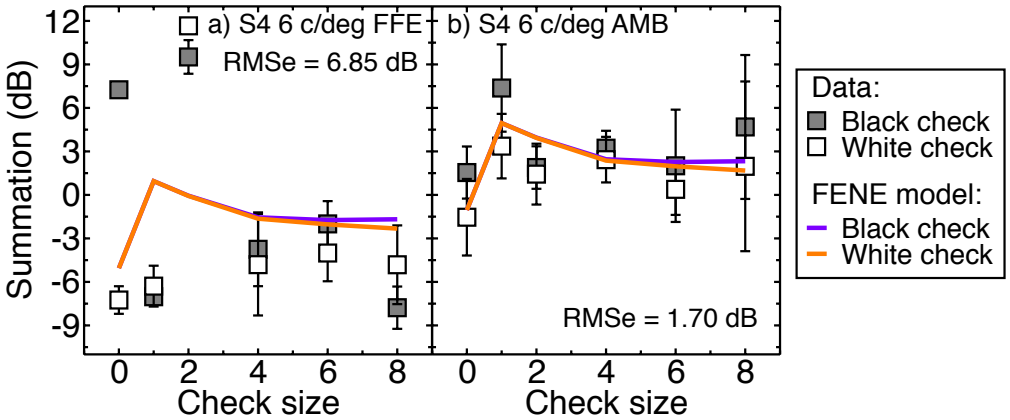

Figure 8.17: Detection thresholds from amblyopic observer S4 for 6 c/deg "Battenberg" stimuli presented to the fellow fixing (FFE) and amblyopic (AMB) eyes. The data are fitted by the prediction from the FENE model.

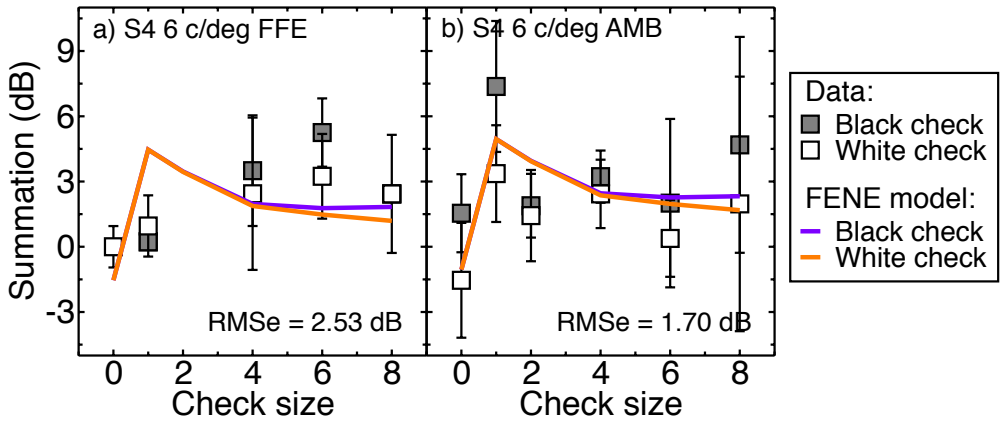

Figure 8.18: Detection thresholds from amblyopic observer S4 for 6 c/deg "Battenberg" stimuli presented to the fellow fixing (FFE) and amblyopic (AMB) eyes. These data are replotted from Figure 8.17 however the three conditions for which thresholds were unusually high in that figure are omitted here. The data are fitted by the prediction from the FENE model.

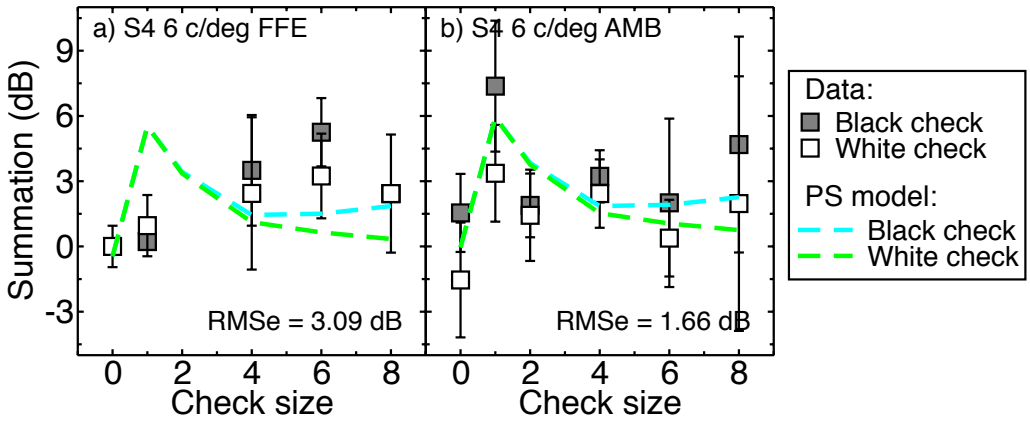

Figure 8.19: Detection thresholds replotted from Figure 8.18. The data are fitted by the prediction from the PS model.

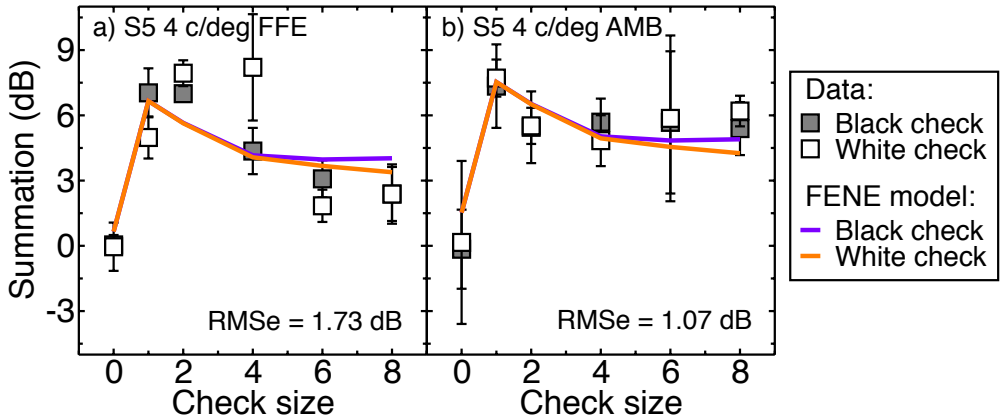

Figure 8.20: Detection thresholds from amblyopic observer S5 for 4 c/deg "Battenberg" stimuli presented to the fellow fixing (FFE) and amblyopic (AMB) eyes. The data are fitted by the prediction from the FENE model.

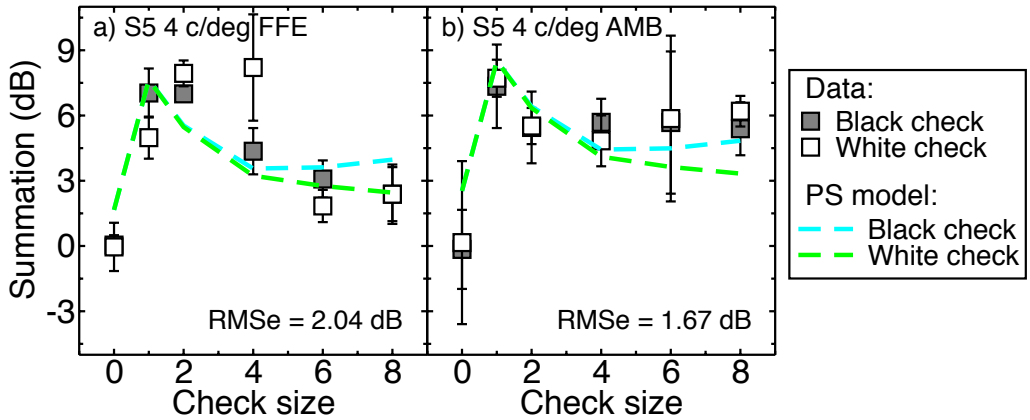

Figure 8.21: Detection thresholds replotted from Figure 8.20. The data are fitted by the prediction from the PS model.

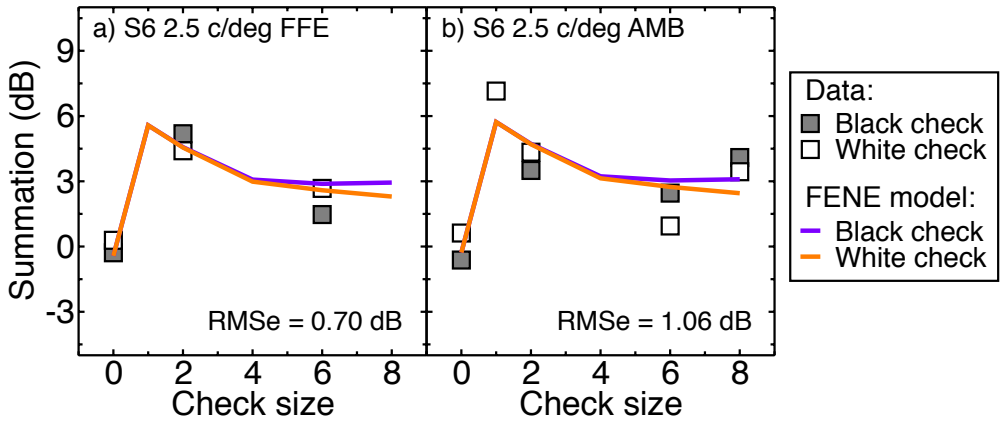

Figure 8.22: Detection thresholds from amblyopic observer S6 for 2.5 c/deg "Battenberg" stimuli presented to the fellow fixing (FFE) and amblyopic (AMB) eyes. The data are fitted by the prediction from the FENE model.

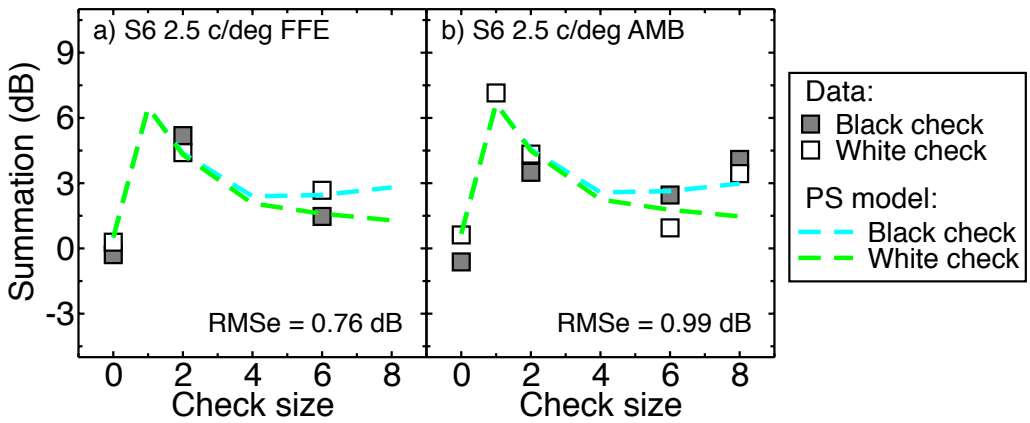

Figure 8.23: Detection thresholds replotted from Figure 8.22. The data are fitted by the prediction from the PS model.

The data collected from the amblyopic observers were quite variable. The quality of the data did however seem to be similar regardless of whether the amblyopic or fellow fixing eyes were being tested. This suggests that the complications of using relatively naïve subjects for this extensive psychophysical study may bear more responsibility for the quality of their data than the pathology of their visual systems. On the other hand, those observers who provided clear data from their fellow fixing eyes (S1, S5, and S6) showed summation in their amblyopic eyes that appeared normal (despite their elevated thresholds), and was better fit in two out of three cases by the FENE model than by the PS model (Figures 8.11b - 8.12b and 8.20b - 8.21b). For observer S6, whose data were slightly better fit by the PS model, very little data had been collected and the difference between the two RMS errors was marginal (Figures 8.22b - 8.23b).

			FENE model		PS model	
Observer	SF (c/deg)	Eye	Figure	RMSe (dB)	Figure	RMSe (dB)
S1	2.5	FFE AMB	8.11	0.88 1.07	8.12	1.16 1.44
S2	2.5	FFE AMB	8.13	1.42 1.44	8.14	2.64 3.13
S3	2.5	FFE AMB	8.15	2.33 4.29	8.16	2.78 4.18
S4	2.5	FFE AMB	8.18	2.53 1.70	8.19	3.09 1.66
S5	2.5	FFE AMB	8.20	1.73 1.07	8.21	2.04 1.67
S6	2.5	FFE AMB	8.22	0.70 1.06	8.23	0.76 0.99

Table 8.3: RMS errors between the data from observer ASB and the FENE and PS model predictions for the fits shown in Figures 8.11 - 8.23 (Figure 8.17 excluded). The rows labelled FFE and AMB refer to the data collected from the observer's fellow fixing eyes and amblyopic eyes respectively.

For the other three amblyopic observers, the data from the amblyopic eyes were more difficult to interpret. Observer S2 shows a loss of the expected initial linear summation peak at 2.5 c/deg (Figures 8.13b - 8.14b), similar to that seen for higher spatial frequencies in the normal observer and the fellow fixing eyes of some of the other amblyopes. Observer S3 shows a greater sensitivity to the 1 by 1 check condition than to the full condition (Figures 8.15b - 8.16b). Both of these results could be explained if the observer was more sensitive to the artefacts introduced by the Battenberg modulation in the checked stimuli than they were to the carrier grating itself (see Section 8.5.3). The data from observer S4 appeared to roughly follow the predicted summation behaviour from the two models, with both the FENE and PS model predictions providing fits of similar quality (Figures 8.18b - 8.19b).

The results of fitting the FENE and PS models to the amblyope data are summarised in Table 8.3. As noted above, in all cases the data from the fellow fixing (FFE) eyes are better fit by

the FENE model (differences in RMSe ranged from 0.06 to 1.22 dB). For the amblyopic (AMB) eyes RMS errors were lower for the FENE model (differences in RMSe ranged from 0.37 to 1.69 dB) in three out of six cases (S1, S2, and S5). For the other three observers the data were only slightly better fit by the PS model (differences in RMSe ranged from 0.04 to 0.11 dB).

8.5.3 Artefacts

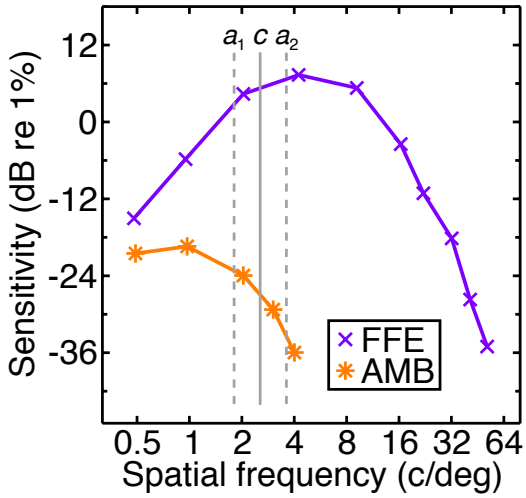

Figure 8.24: Example contrast sensitivity function (CSF) from the fellow fixing (FFE) and amblyopic (AMB) eyes of an amblyope replotted from Hess et al. (1978), Figure 7. The sold grey line labelled c marks a possible stimulus carrier frequency (2.5 c/deg) and the dashed lines labelled a_1 and a_2 show the frequencies of the artefacts introduced below and above that spatial frequency by the 1 by 1 check Battenberg modulation.

The loss of the initial 6 dB peak from short-range linear summation within receptive fields cannot be explained within the single-channel models tested here. Instead, an explanation for this effect must rely on interactions from other spatial frequency and orientation channels (see Section 2.3.2). As detailed in Section 3.4.6, Battenberg stimuli are created by multiplying several component images together. These are a sine-wave carrier grating and a rectified orthogonal grating at half the spatial frequency, which are multiplied to generate the "full" stimulus (see left column of Figure 8.25). The checked Battenberg stimuli are then generated by multiplying this full stimulus by a square-wave checkerboard (see left column of Figure 8.26).

The Fourier transforms of the images used at each stage are shown in the right columns of Figures 8.25 and 8.26. As multiplication in the spatial domain is equivalent to convolution in the Fourier domain, the Battenberg stimuli acquire artefacts (easily visible in the Fourier domain) at spatial frequencies and orientations other than that of the carrier. Those which may have a confounding effect in this study are the ones introduced by the squarewave modulating plaid used to create checkerboard pattern in the Battenberg stimuli. For the 1 by 1 check stimuli, this manipulation introduces stimulus energy at spatial frequencies at factors of $\sqrt{2}$ above and below the carrier. When the carrier spatial frequency is not at the peak of the CSF (i.e. is

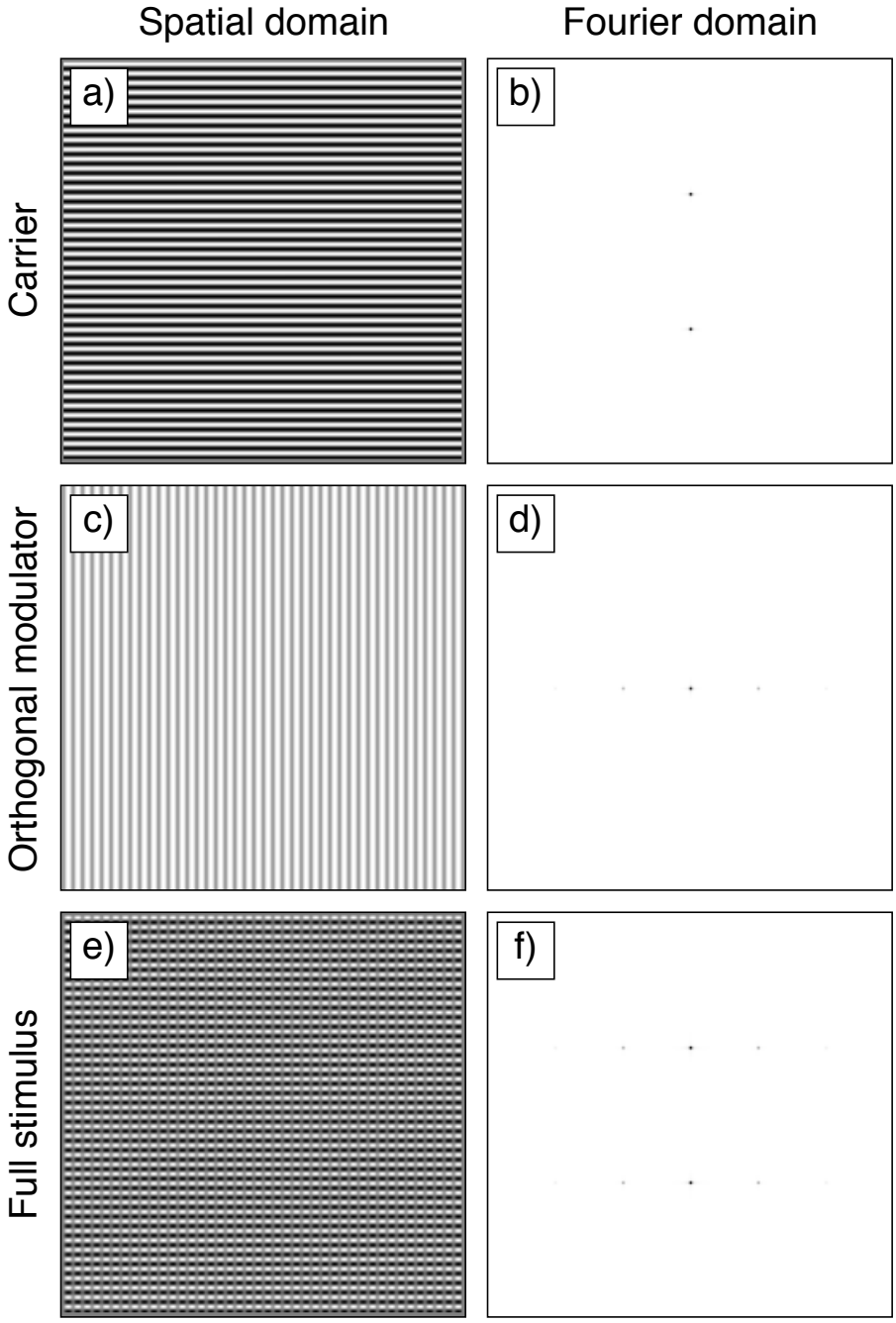

Figure 8.25: Images of stimulus components (left column) and Fourier transforms of those images (right column). In the spatial domain, the full stimulus (bottom row) is the product of the carrier grating (top row) and the rectified orthogonal modulator grating (middle row). The modulator is generated at half the spatial frequency of the carrier (therefore it has the same periodicity after full-wave rectification).

not among the most detectable frequencies) it is possible that these artefacts introduced at other, possibly more detectable, spatial frequencies will lower detection thresholds (see Figure 8.24). Such an effect would explain the absence of linear summation for the smallest check size in several observers. This could be the case whether those data are from the amblyopic eye at a lower spatial frequency, or a normal eye at a higher spatial frequency.

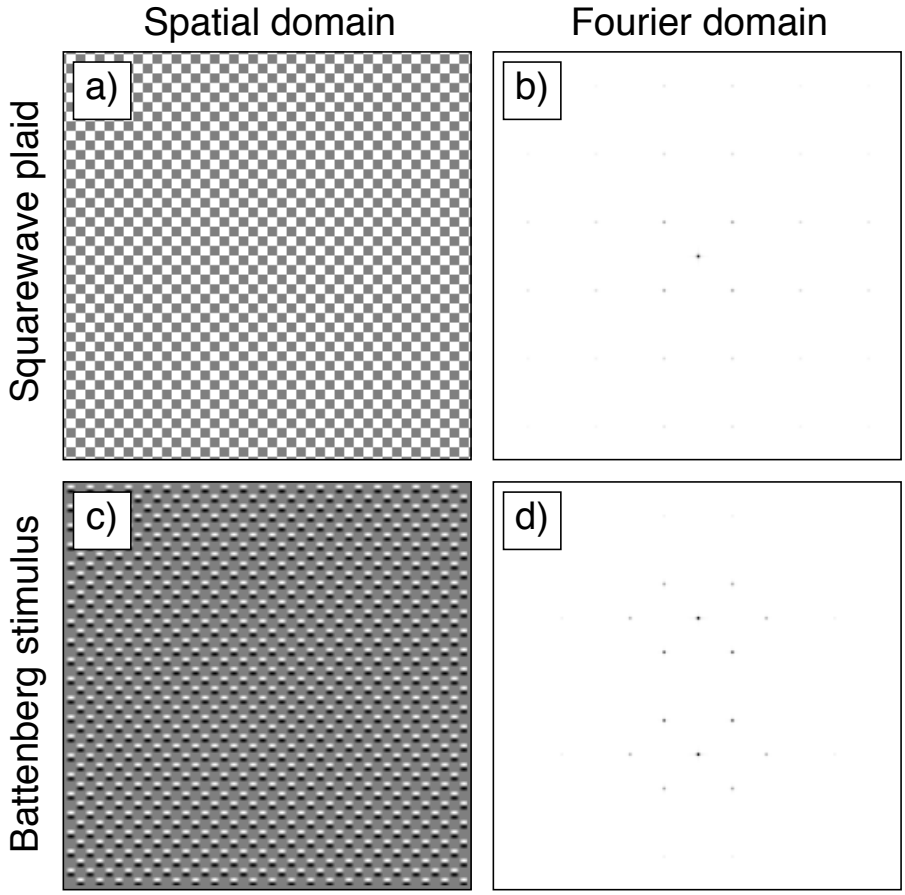

Figure 8.26: Images of squarewave plaid modulator and 1 by 1 check "Battenberg" stimulus (left column) and Fourier transforms of those images (right column). The Battenberg stimulus (bottom row) is the product of the full stimulus from Figure 8.25e and the squarewave plaid modulator (in the spatial domain).

Figure 8.3 shows a portion of the CSF for observer ASB, with the artefact frequencies for the 4 c/deg and 8 c/deg stimuli labelled a_2 and a_3 respectively. The figure shows that these frequencies are more detectable than the carrier in each instance, supporting the possibility that the observer may have improved their performance for the 1 by 1 check stimuli by detecting these artefacts. This would explain the reduced short-range summation found for these conditions (see Figures 8.6 and 8.8). For the 2.5 c/deg stimuli there was no data from the relevant frequency (1.8 c/deg) to determine whether this stimulus was at the peak of the CSF, therefore it is not known whether an explanation based on sensitivity to artefact frequencies would predict the *presence* of the full linear summation effect for stimuli of this spatial frequency.

Amblyopic observer S1 showed full linear summation for the 1 by 1 check stimuli in both eyes. Observer S2 showed slightly less than linear summation in the fellow fixing eye, and no summation in the amblyopic eye. Figure 8.27 shows contrast detection thresholds collected from these two amblyopic observers in their fellow fixing and amblyopic eyes for gratings at the carrier spatial frequency (2.5 c/deg) and at a factor of $\sqrt{2}$ below the carrier frequency (1.8 c/deg). The stimuli used were 42 by 42 cycle horizontal gratings generated with a fixed num-

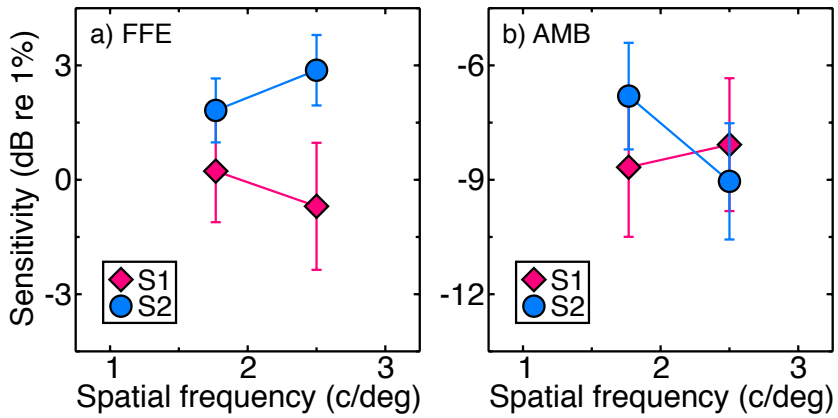

Figure 8.27: Contrast detection thresholds for two amblyopic observers (S1 and S2) at 1.8 and 2.5 c/deg. Data are shown for the fellow fixing (FFE) and amblyopic (AMB) eyes.

ber of cycles on the screen. The methods used were otherwise similar to those presented in the methods section of this chapter.

For the observer who shows normal linear summation (S1), thresholds for the artefact spatial frequency were 1.05 dB higher than those for the carrier spatial frequency in the amblyopic eye (Figure 8.27b). For the observer who does not show any summation for the 1 by 1 check stimuli (S2), thresholds were 2.24 dB lower for the artefact spatial frequency than for the carrier spatial frequency. In the fellow fixing eyes (Figure 8.27a) the differences were smaller direction of these effects was reversed, with observer S1 showing a 0.92 dB advantage for the artefact spatial frequency, and observer S2 showing a 0.59 dB disadvantage. These findings support the possibility that the lack of summation for the 1 by 1 check stimuli in the amblyopic eye of observer S2 is due to the shape of the CSF for this observer favouring the detection of the artefacts introduced to the stimulus by the Battenberg modulation, however a full explanation would require a model of how stimulus information is combined across different spatial frequency channels.

8.6 Discussion

8.6.1 Area summation in amblyopes appears to be normal

The results from this study support the established finding that although contrast detection thresholds are higher in amblyopia, the summation of contrast over area occurs in a similar manner to that found in normal observers (Hess & Howell, 1978; Hess & Campbell, 1980). For three out of six observers (S1, S2, and S5) the amount of summation found was greater than that predicted by a probability summation model, instead favouring a linear sum of the squared local filter outputs. This behaviour would be consistent with the noisy energy model put forward in previous chapters, if the observer were unable to match a template to the

checkerboard pattern in the stimuli here and instead summed signal and noise over the stimulus extent. For the other three observers (S3, S4, and S6) the results were inconclusive as similar RMS errors were obtained by fitting either model prediction (probability summation or noisy energy) to the data.

8.6.2 Complications involved in the use of the Battenberg stimulus

The results from both a subset of the amblyopic observers and the normal observer at higher spatial frequencies are fundamentally incompatible with the single-channel models used here. One possible explanation of these results would be that the artefacts introduced at other spatial frequencies in the Battenberg stimulus generation were contributing to the detection of the checked stimuli. Where larger amounts of summation are found (such as in the observers that are best fit by the FENE model prediction here) this is no cause for concern, as a contribution to detection from the artefacts can only *decrease* the measured level of summation. Where less summation is found than expected however these artefacts may be responsible. The possibility that unintended artefacts may affect the results obtained using Battenberg stimuli should be considered in studies that use them.

CHAPTER 9

Summation of orientation signals

9.1 Motivation and summary

Previous chapters in this thesis have focused on how threshold contrast signals are combined across the visual field. In this chapter, the same techniques used to study those processes are applied to an investigation of how local orientations in a suprathreshold pattern are pooled over space into a coherent global percept. Although previous studies have addressed the signal combination process at length (i.e. how an estimate of global orientation is calculated from the local samples), the strategy by which samples are chosen to be pooled (e.g. whether the observer is able to ignore samples from irrelevant locations) is a topic that has attracted surprisingly little attention previously. A set of orientation discrimination experiments are performed here with novel "Battenberg" stimuli. The results suggest the combination of orientation signals over space to be a two-stage process. The first stage is a mandatory integration of local signals over a fixed area, limited by internal noise at each location. The second stage is a task-dependent flexible combination of the outputs from the first stage.

9.2 Introduction

9.2.1 Combining orientation signals over space

The perception of coherent textures requires the integration of orientation signals over space. The definition of regions in an image that "belong" to the same texture is a necessary intermediate step to higher-level processes such as finding boundaries between different textures (Marr, 1982; Landy & Graham, 2004). Although texture boundaries are not the only cue to used to segregate objects from each other in normal viewing they do provide at least *some* information (see Figure 9.1). Research on human behaviour in this area has focused on address-

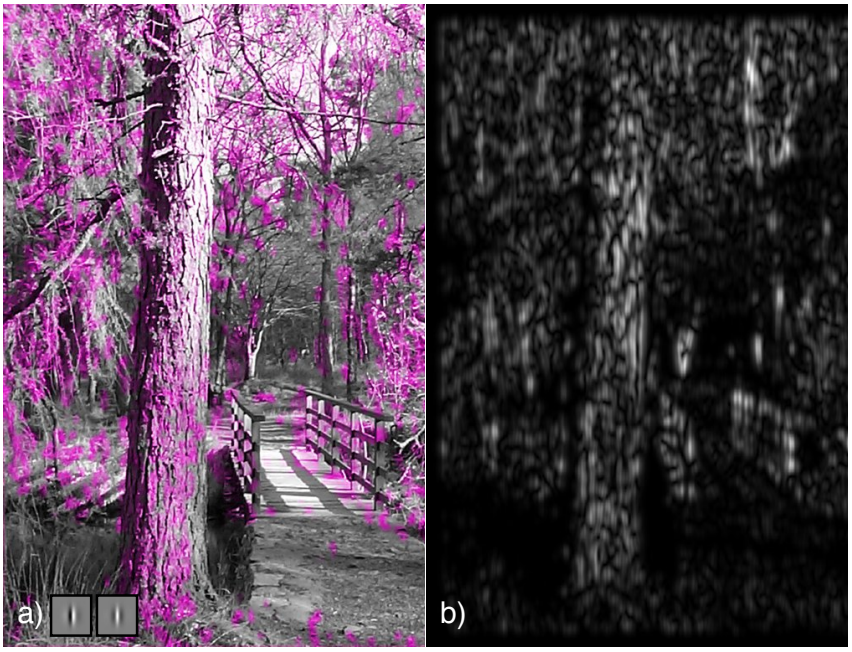

Figure 9.1: Demonstration of coherent texture. Panel a) shows an image of a natural scene. This was filtered by a pair of quadrature-phase vertical log-Gabor elements (inset), the complex response at each location was calculated as the Pythagorean sum of the sin and cosine responses. Pink locations in the image indicate where this response exceeded an arbitrary threshold. Panel b) shows the vertical energy in the image at the spatial scale at which the filtering in panel a) was conducted. The tree trunk on the left of the image, and the fences on each side of the bridge, are represented by clear regions of coherent vertical texture.

ing the process by which signals are combined, rather than the effects of the spatial arrangement of the signals in the stimulus. Differential behaviours in psychophysical tasks would result from the various possible strategies by which signals could be pooled into the combination process. It is recognised that the pooling strategy leading to the perception of a coherent texture must involve both a local integration phase* where the orientation statistics at each location are estimated, and further operations performed over those local estimates (Vorhees & Poggio, 1988; Sagi, 1990; Dakin & Watt, 1997). This study focuses on examining the strategies used for pooling samples over space in the two stages by measuring and modelling the integration of local orientation signals across the visual field.

9.2.2 Signal combination processes

Effects of spatial configuration aside, the process by which the visual system calculates a global orientation from an array of local orientations has been found to depend on the task set to the

*The structure of the local mandatory phase is described by a linear-nonlinear-linear (LNL) model (Landy & Graham, 2004), where the non-linear stage sandwiched between the two linear signal combination stages is necessary for the predictions to be different from those made by a simpler linear model (e.g. for stimuli where the local responses from the first linear stage are out of phase).

observer. Similar dependencies have also been reported in studies that investigated the integration of local motion signals. When the observer is required to distinguish between stimuli with weak horizontal or vertical orientation coherence (i.e. with a large difference between the two target orientations), observers filter the image at the two potential target orientations and then choose the orientation of the more activated filter (Husk et al., 2012). A winner-takes-all process similar to this has been found in analagous studies performed with motion stimuli in monkeys (Salzman & Newsome, 1994).

Under conditions where finer judgements of the global orientation of a texture need to be made, the observer calculates the vector average of the local orientations (Dakin & Watt, 1997; Webb et al., 2010). Similar changes in the combination process used by observers based on the difference between the discriminated orientations have been demonstrated in the motion domain (Nichols & Newsome, 2002; Webb et al., 2007). The ideal strategy for any task would be to use all of the information available from the stimulus on each trial to calculate which response is most likely to be correct (maximum likelihood estimation, or MLE). Studies designed to produce distinct predictions from MLE have found that observers use this process when the stimulus is displayed for a sufficient duration (longer than 1 second in Webb et al., 2010).

The large orientation differences and short stimulus durations used in the current study would be expected to cause the observer to max over filter outputs (the design of this study is similar to that of Husk et al., 2012), however it is not necessary to assume that the observer makes use of a particular process for the purposes of this study (see Section 9.5.2). Instead, what is of interest is the effect of the pooling strategy used to select the samples for this combination. The filter maxing and vector averaging models are both implemented here in order to demonstrate that they make near-identical predictions for the tasks in this study.

9.2.3 Pooling strategies and summation effects

Most signal combination processes would predict an improvement in performance for detecting weak signals as the number of samples increases. Provided that the noise affecting each sample is at least partially independent, the limiting effect of the noise on performance can be reduced by exploiting the information from multiple samples. This will improve performance whether the observer is filter-maxing, vector-averaging, or performing MLE. Different strategies for pooling signals over space make different predictions for how performance should improve with additional signal samples. Previous work where the number of samples available for combination is varied have found conflicting results of either improvements reflecting ideal summation under a flexible pooling strategy (Jones et al., 2003), or no benefit from increasing the number of samples whatsoever (Husk et al., 2012). Dakin (2001) found "entirely flexible" combination with respect to signal location over a limited number of samples in the display. This was presented as an "information limit" for orientation integration.

9.2.4 This study

Although several studies have investigated the combination process underlying the judgement of global orientation, relatively little attention has been paid to the pooling strategy that is used to select the samples for combination. The summation effects resulting from increasing the number of samples available for integration are investigated here using novel Battenberg stimuli and computer modelling. The Battenberg stimuli allow for manipulation of the spatial arrangement of signal within a stimulus of fixed extent and eccentricity. This reduces the confounding effects of any inhomogeneities in sensitivity for performing the global orientation task. Jones et al. (2003) suggested that such an effect may have reduced the level of summation measured in their study. The results of the experiments conducted in this chapter show that the pooling of orientation information is a two-stage process, with mandatory local integration (affected by internal noise at each location) followed by flexible long-range pooling.

9.3 Methods

9.3.1 Equipment

Stimuli were presented on a gamma-corrected CRT monitor using Psychtoolbox (Brainard, 1997; Kleiner et al., 2007) running under MATLAB. The data collection for these experiments was split between two different equipment setups. The first was an Apple Macbook Pro with an NVIDIA GeForce 9600M graphics card presenting stimuli on a Philips MGD403 monitor, the second was an Apple Macbook Pro with an NVIDIA GeForce 8600M graphics card presenting stimuli on a Compaq monitor. The monitors had refresh rates of 75 Hz and 90 Hz, and mean luminances of 77.2 and 26.9 cd/m^2 respectively. Observers viewed the monitors from a distance of 0.51 metres. At this viewing distance the screen had a resolution of 24 pixels per degree of visual angle, giving 6 pixels per cycle for the 4 c/deg stimuli used here.

9.3.2 Stimuli

Stimuli were 29 by 29 element arrays of 4 c/deg cosine-phase log-Gabor patches (spatial frequency and orientation bandwidths of 1.6 octaves and $\pm 25^\circ$ respectively; see Section 3.4.5), spaced 1 degree apart in a square grid. Stimuli were displayed at 80% delta-contrast. Each log-Gabor was either a potential signal or a noise element. Potential signal elements had probability $P(\text{coherence})$ of assuming the target orientation (which was either horizontal or vertical), otherwise they assumed an orientation drawn at random from a uniform distribution. All noise elements assumed random orientations. The orientation of each element was rounded to the nearest degree before stimulus generation. The range of potential element orientations was 0° to 179° (angles beyond this were redundant for the cosine-phase elements used here).

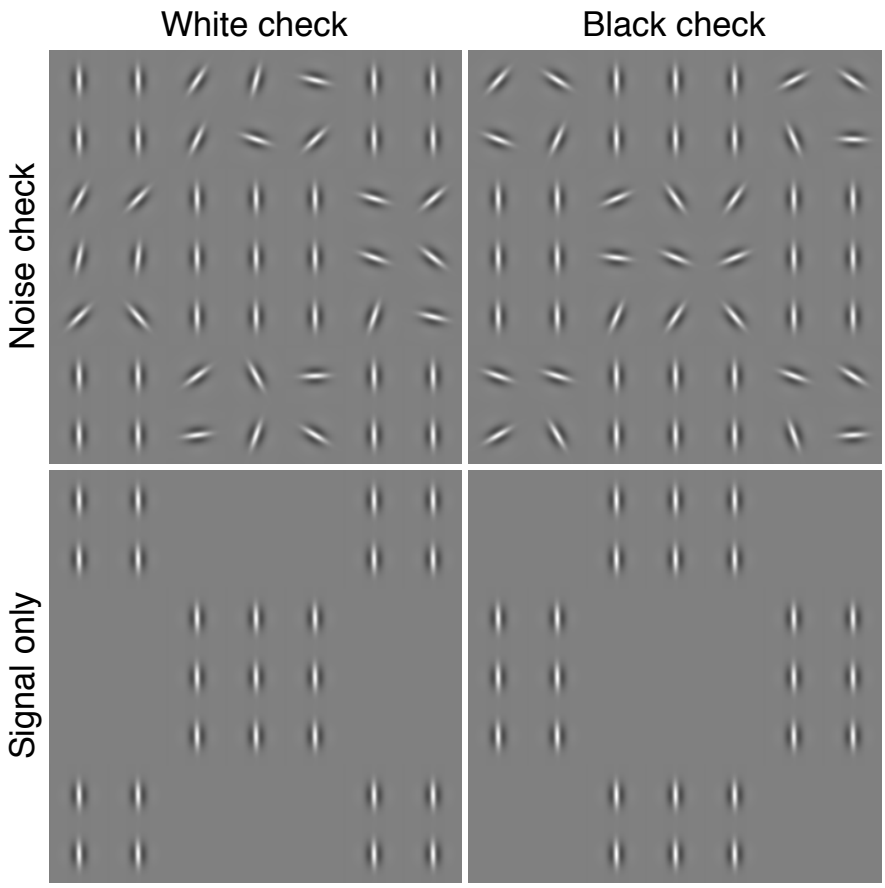

Figure 9.2: Example of the stimulus design used in these experiments. The stimuli shown are 7 by 7 element arrays (smaller than the 29 by 29 arrays used in this study) with a 3 by 3 check size. The figure includes "white" and "black" check versions of the noise check and signal only stimuli, shown at 100% coherence.

Two types of stimuli were tested, "full" and "Battenberg". In the full stimuli all of the elements were potential signal elements. For the Battenberg stimuli, the potential signal elements were assigned to locations in the stimulus defined by a checkerboard (a square-wave plaid). This gave a stimulus tiled with square signal and non-signal regions. Two types of Battenberg stimuli were tested. For the "noise check" condition the non-signal regions contained randomly-oriented elements. For the "signal only" condition the non-signal regions were blank (mean luminance). Each of the Battenberg conditions contained the same total number of elements, but approximately half as many signal elements as the full (all-signal) condition (see Table 9.1).

The spatial arrangement of the signal regions in the stimulus was manipulated by adjusting the frequency and the phase of the square-wave plaid modulator† that defined the checkerboard. Decreasing or increasing the frequency made the signal regions larger or smaller respectively, and this was used to create the five different check sizes. These gave stimuli tiled with 1, 3, 5, 9 and 15 element square signal regions (i.e. the largest check size had 15 by 15 element

†This modulator was generated as the product of vertical and horizontal square wave gratings (scaled 0 - 1), rather than by summing orthogonal oblique gratings.

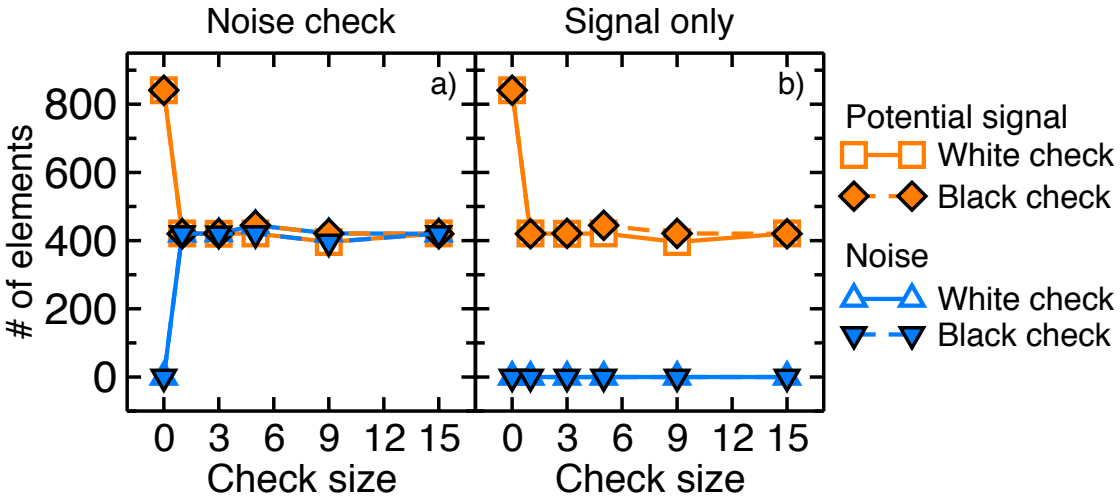

Figure 9.3: The numbers of potential signal and noise (always randomly-oriented) elements in each stimulus. In the noise check stimuli (panel a) the total number of elements (841) remained constant. In the signal only stimuli (panel b) the checked stimuli contained approximately half as many elements as the full (0) stimulus.

	White check		Black check	
Check size	# signal elements	Proportion	# signal elements	Proportion
1 by 1	421	50.1%	420	49.9%
3 by 3	420	49.9%	421	50.1%
5 by 5	421	50.1%	445	52.9%
9 by 9	396	47.1%	421	50.1%
15 by 15	421	50.1%	420	49.9%

Table 9.1: Numbers and proportions of potential signal elements in the various "Battenberg" stimuli used in this study. The total number of elements in the full Battenberg stimulus was 841. Noise check stimuli always contained 841 elements, with the non-signal elements set to random orientations. Signal only stimuli did not contain any elements other than those which were potential signal elements.

"checks"). The phase of the modulation was also manipulated to test stimuli in both cosine ("white check") and anticosine ("black check") phase. Miniature example stimuli are shown in Figure 9.2. The numbers of potential signal and noise (always randomly oriented) elements in each stimulus are plotted in Figure 9.3. The full stimuli were referred to as having a check size of zero for the purposes of plotting the data.

9.3.3 Observers

Seven observers were used. One was the author (ASB), four were experienced psychophysical observers (DHB, JSH, RJS & SAW), and two were naïve undergraduate students (LFE & VRP). All had either normal or corrected-to-normal vision.

9.3.4 Procedures

A blocked single-interval identification task was performed to find the threshold coherence for each check size, phase (white vs. black check), and Battenberg type (noise check vs. signal only). Threshold coherence was tracked using a pair of three-down one-up staircases (maximum 120 trials or 12 reversals), one for horizontal signal trials and the other for vertical signal trials (see Section 3.5). The staircases for the two signal orientations were interleaved randomly. Once the staircase for one orientation had terminated, dummy trials (where no data were recorded) were still presented with that orientation until the staircase for the other orientation terminated. Staircases started at a high level in order to inform the observers of what stimulus to expect in each block.

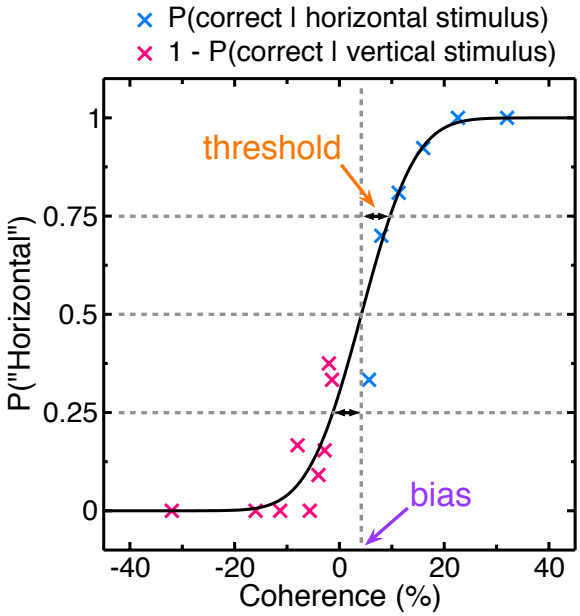

Figure 9.4: Showing the method by which threshold coherence levels were calculated. Data from the two interleaved staircases (one for each signal orientation) were combined and fitted with a cumulative normal to give a single psychometric function.

Stimuli were presented for 250 ms. Stimulus onset was accompanied by a beep. Observers fixated centrally, with the help of a black fixation dot that was shown between trials. A keyboard was used for the response. The observers pressed a key to indicate whether the stimulus contained either "Horizontal" or "Vertical" coherence. The observers' response was followed by a feedback beep that indicated whether it was correct or incorrect, and then a 300 ms pause before the presentation of the next stimulus. Each observer performed four repetitions for each combination of check size (0 to 15), Battenberg modulator phase (white or black check), and Battenberg type (noise check or signal only). As the full (0) stimulus was identical regardless of Battenberg type or modulator phase, each observer collected four times as much data for this condition (sixteen repetitions). These were averaged to give a single threshold per observer.

9.3.5 Analysis

Data from the horizontal and vertical staircases were combined into a single psychometric function for each repetition, condition, and observer. This was then fit by a cumulative normal function using Palamedes (see Section 3.6; Prins & Kingdom, 2009). The fitted function gave the probability of responding "Horizontal" to either a vertical stimulus (plotted as negative coherence) or a horizontal stimulus (plotted as positive coherence). The coherence level at which the function reached P("Horizontal") = 0.5 gave the bias for the observer categorising a stimulus as horizontal rather than vertical (or vice-versa), and the threshold could be calculated as the difference between the coherence level at this point and that where P("Horizontal") = 0.75 or P("Horizontal") = 0.25.

9.4 Results

9.4.1 Noise check

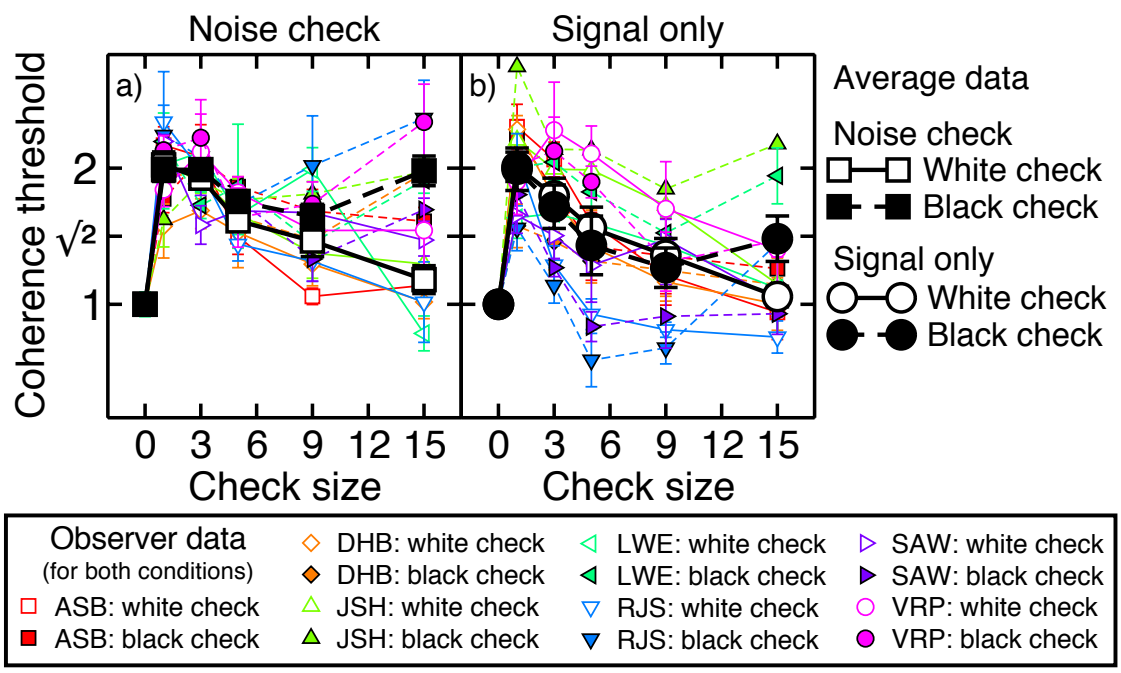

Figure 9.5: Results from individual observers (see legend below) and the average results across observers (right). Performance is expressed as the threshold proportion of coherent elements in the signal regions. Thresholds are expressed as multiples relative to that for the full (0) condition (these data are replotted in Figure 9.6). Results are shown for the noise check (a) and signal only (b) conditions. Error bars show ± 1 standard error here and in all other graphs.

Results from the seven observers, and the average of their results, are shown in Figure 9.5. Coherence thresholds (proportion of elements in the potential signal region that assume the

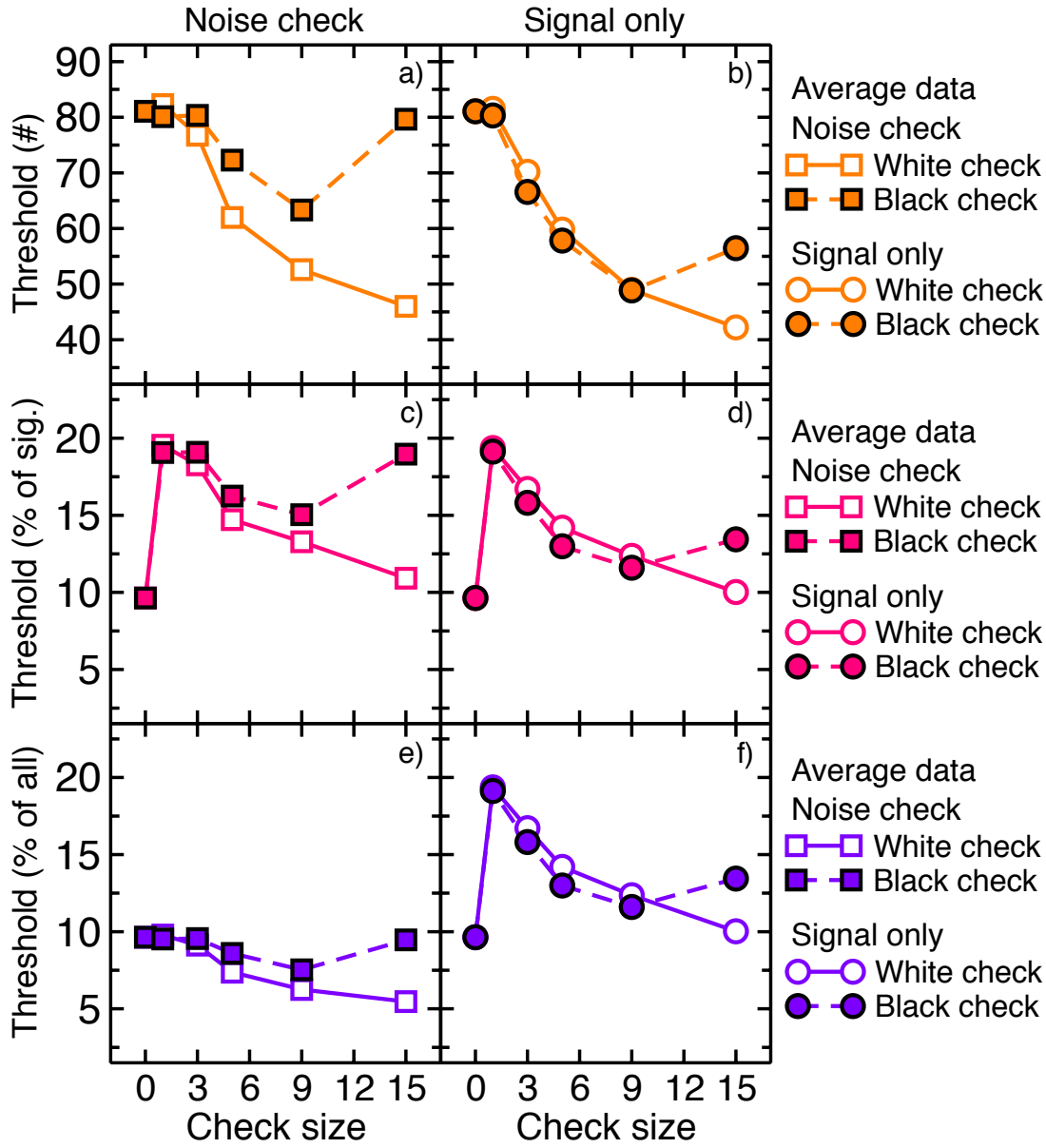

Figure 9.6: The data from Figure 9.5 are replotted here as the threshold number of elements in the stimulus (panels a-b), the threshold percentage of signal elements in the potential signal region (panels c-d), and the threshold percentage of signal elements in the entire stimulus (panels e-f). Other figures in this chapter use the representation of the threshold signal level shown in panels c) and d), with these "coherence thresholds" expressed as multiples relative to that of the full (0) stimulus.

target orientation) for each stimulus are plotted as multiples relative to that of the full (0) condition, which had a threshold of approximately 10%. Figure 9.6 provides more information on how to interpret this figure. Figure 9.5a shows the results for the noise check condition. For smaller check sizes (1 to 3) the threshold coherence for the Battenberg stimuli was approximately double that for the full (0) stimulus. This means that approximately the same number of signal elements across the entire stimulus were required to reach the threshold performance level in each case (see Figure 9.6a). For medium sizes (5 to 9) the threshold elevation

decreased to $\sqrt{2}$, which is consistent with a strategy that uses information from potential signal regions but ignores irrelevant (noise-only) regions (as the noise resulting from the combination of multiple equally-noisy samples is proportional to the square root of the number of samples combined). This point will be returned to in the modelling.

One of the purposes of the Battenberg stimulus design is to reduce the effect of visual field inhomogeneities in sensitivity on the measurement of area summation. For stimuli with large check sizes relative to their extent though, these effects will return. For the largest (15) check size, performance diverged dependent on whether the stimulus was in the white check (foveal signal, peripheral noise) or black check (foveal noise, peripheral signal) phase. Thresholds were almost as low for the white check stimulus as they were for the full (0) stimulus. This means that the observers required around half as many signal elements across the entire stimulus when most of those elements were presented in the centre of the display (Figure 9.6a). For the black check stimulus, coherence thresholds were approximately double that of the full (0) stimulus. This means observers required the same number of signal elements in the largest black check stimulus as they did in the full (0) stimulus (Figure 9.6a). This would be consistent either with a relative insensitivity for this task in the periphery or a failure in segregating the noise present in the centre of the display (this is discussed further below).

9.4.2 Signal only

Figure 9.5b shows the results for the signal only condition. For the smaller check sizes in the signal only condition, thresholds increase to approximately double that for the full (0) stimulus in a similar manner to that seen in the noise check condition. This is unexpected, as the predicted threshold for this condition based on the stimulus properties alone would be a factor of $\sqrt{2}$ above that for the full (0) condition (there are no noise-only elements, so the threshold would be proportional to the square root of the number of elements in the display). These results suggest that the limiting noise for this task does not come from the responses to the individual elements (i.e. is not "early"). The noise cannot be "late" and constant across conditions however, as this would predict the same performance level for all of the checked stimuli. Instead, these results suggest that observers are mandatorily integrating internal noise from blank display regions (or are limited by late noise that is proportional to the monitored area) for the smaller check sizes but are able to exclude this noise for the larger check sizes. A possible explanation for this behaviour will be offered in the modelling (Section 9.5.7).

For the largest (15) white check stimuli averaged thresholds are similar in the noise check and signal only conditions, suggesting that the pooling of samples (and segregation of noise) in these two conditions is similar. Performance for the largest (15) black check stimulus in the signal only condition is a factor of $\sqrt{2}$ better than that in the noise check condition, meaning that at least part of the deficit for detecting that stimulus was due to inefficient pooling of noise in the noise check condition rather than a relative insensitivity for performing the task

in the periphery (e.g. it may be more difficult for the observer to keep track of the signal and non-signal regions in the noise check condition). Observers JSH and LWE show similar performance for the largest black check stimulus in the two conditions (noise check and signal only), suggesting that in their case it *is* the insensitivity of the periphery that is the limiting factor. Excluding these observers from the averaging (not shown) does not affect the thresholds for the largest (15) check sizes in the noise check condition, but does lower the threshold for the black check stimulus in the signal-only condition to approximately $\sqrt[4]{2}$ (increasing the difference between the noise check and signal only conditions for the largest black check stimulus). Possible explanations for the advantage found for the largest white check stimuli are addressed in the modelling section.

9.5 Modelling

9.5.1 Monte Carlo simulations

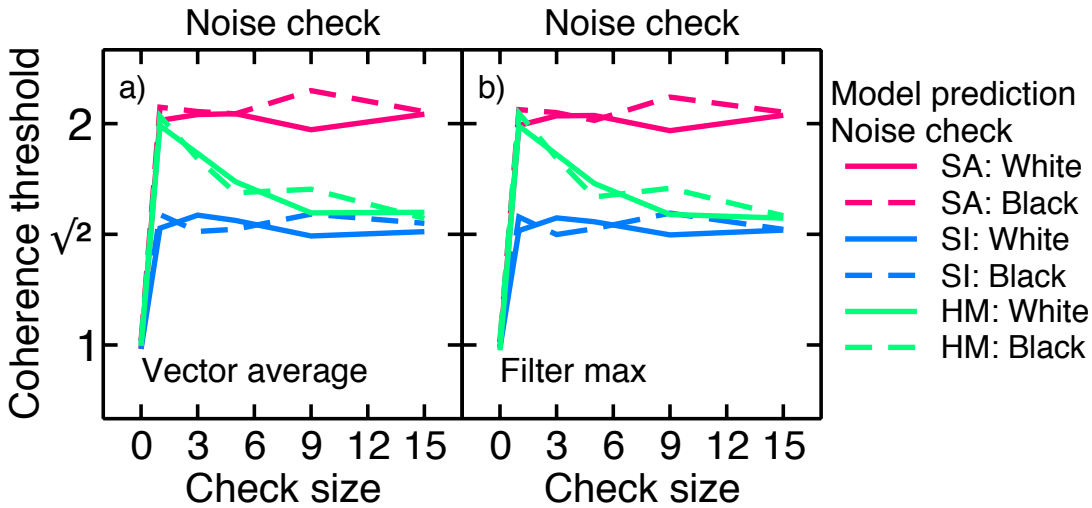

Figure 9.7: Predictions for the noise check condition from models using the vector averaging (a) and filter maxing (b) combination processes and a variety of different pooling strategies. Each plot shows predictions for models that combine over all elements (SA), combine ideally (SI), and a two-stage hybrid model (HM) of mandatory local combination followed by ideal pooling (see Section 9.5.5).

A set of models was developed in order to investigate the pooling strategy used by the observers to perform the task in this study. The main aim of the modelling was to explain the surprising equivalence between the smallest check sizes in the noise check and signal only conditions. Models were developed using stochastic Monte Carlo methods (Section 5.1). A set of model observers was developed with different pooling strategies and combination processes in MATLAB and run through 2,000 simulated trials per stimulus level of a method of constant stimuli (MCS) version of the experiment. The thresholds for each model observer

performing each condition were then calculated using the simulated data and expressed as relative thresholds (i.e. the threshold for each Battenberg condition expressed as a multiple of that for the full (0) condition) allowing them to be compared directly to the data (obviating the need to fit the models to the human data).

9.5.2 Combination processes

Several combination processes for the calculation of global orientation from individual local samples have been suggested previously. Here I implement the two strongest candidates that the observers may be using for the tasks in this study: vector-averaging of individual element orientations (found for orientation discrimination of glass patterns by Webb et al., 2010) or selecting the orientation of the most strongly activated filter (found for an orientation coherence task similar to this one by Husk et al., 2012). These two combination processes make very similar predictions for how performance on the task investigated here should depend on the number of integrated signal and noise samples (see Figure 9.7). To demonstrate this however, and for the sake of completeness, both are implemented here.

9.5.3 Vector averaging

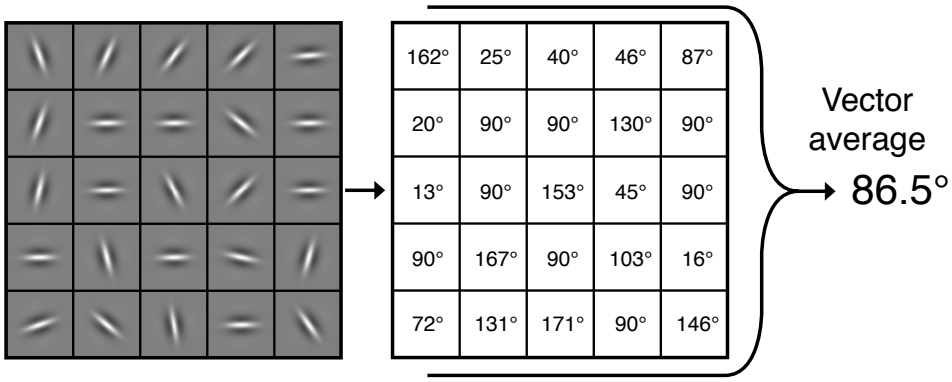

Figure 9.8: Demonstration of the vector averaging combination process. The orientation of each pooled element is input into the vector averaging formula (see Equation 9.1), from which the average orientation is then calculated.

In the vector averaging model (Dakin & Watt, 1997), each pooled element is represented as a vector with magnitude m_i and orientation θ_i (see Figure 9.8). It is assumed that the observer is able to extract the orientation of each element (in the model this is implemented by taking the orientations directly from the stimulus generation procedure). These are then combined using vector averaging to get the average orientation

$$\theta_{\text{avg}} = \frac{1}{2}\tan^{-1}\left(\frac{\sum_{i=1}^{n} m_i \sin 2\theta_i}{\sum_{i=1}^{n} m_i \cos 2\theta_i}\right). \tag{9.1}$$

Note that the local orientations are doubled before averaging, and that the output of the vector averaging operation is halved. This wraps the orientations at 180° (rather than at 360°, which is the usual limit) because each element in the display is symmetrical across its major and minor axes (a 90° element is identical to a 270° element).

When every element is weighted equally all elements are represented by unit vectors ($m_i = 1$). In the case where the elements have different expected signal-to-noise ratios (e.g. the second stage of the two-stage hybrid model below), the magnitudes of the local vectors are each weighted by a template in order to control the contribution each local vector makes to the calculated average (see Section 9.5.5). The model then picks the potential target orientation closest to the calculated average orientation. The predictions from vector averaging models using the different pooling strategies given in Section 9.5.5 are shown in panel a) of Figure 9.7.

9.5.4 Filter maxing

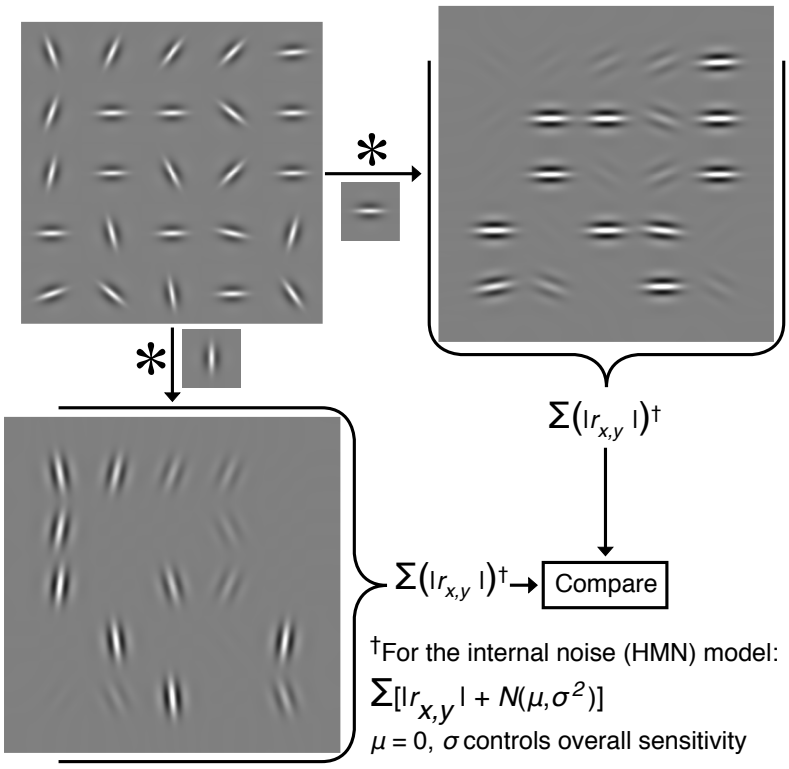

Figure 9.9: The filter maxing combination process. The stimulus is filtered at the two possible target orientations by log-Gabor patches matched to the signal elements. The magnitudes of the filter responses at each pooled location are then summed over the image and the responses from the two orientations are compared to pick the larger.

In the filter maxing model (Jones et al., 2003), the stimulus is first filtered at the two potential target orientations (see Figure 9.9). The filter elements were a pair of log-Gabor elements

with the same properties as those used in the generation of the stimuli. The rectified filter responses in the pooled regions are then summed over the image for each orientation and these values are compared to each other. The observer picks the orientation with the largest filter response. The predictions from filter maxing models using the different pooling strategies given in Section 9.5.5 are shown in panel b) of Figure 9.7.

Comparing the model predictions for the three different pooling strategies between the two panels in Figure 9.7 shows that the combination process used does not affect the model prediction. This is unsurprising as previous studies that have found differential predictions from these models have done so by varying the distribution of the element orientations (e.g. Webb et al., 2007; Husk et al., 2012), whereas this is kept constant across the conditions tested here. Due to the similarity of the predictions from the two combination processes, only the predictions from the filter maxing models are used in the rest of this chapter.

9.5.5 Pooling strategies

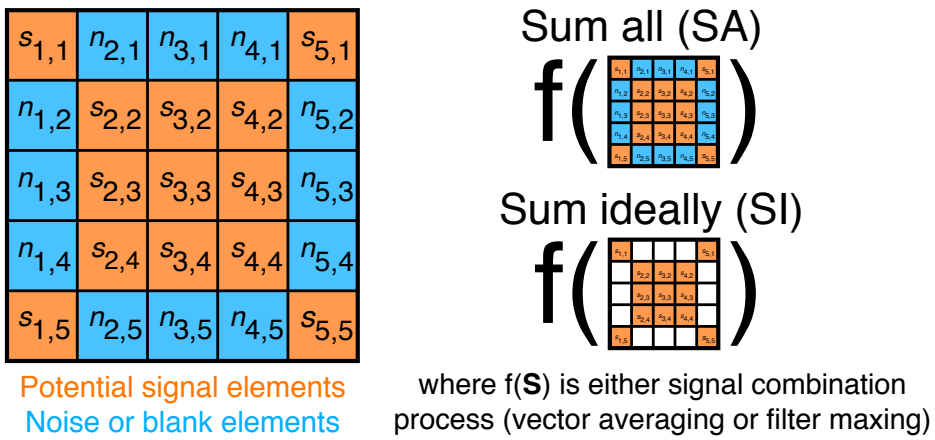

Figure 9.10: Showing the "sum all" (SA) and "sum ideally" (SI) pooling strategies. The orange regions in the diagram show locations of potential signal elements, the blue regions show locations of noise or blank elements. The SA strategy combines information from all locations in the stimulus (signal and noise/blank), the SI strategy uses only information from the potential signal regions. The function $\mathrm{f}(\mathbf{X})$ represents either vector averaging or filter maxing depending on which combination process is being used.

The simplest pooling strategy considered is the "sum all" (SA) strategy, where the observer combines information from every element in the stimulus regardless of whether it is a noise or potential signal element (see Figure 9.10). Models with this strategy predict that the threshold coherence for the noise check stimuli should be approximately twice that for the full (0) stimulus (Figure 9.7), which is equivalent to requiring the same threshold proportion of coherent signal elements across the whole stimulus ("approximately" because the checked stimuli do not contain *exactly* half as many potential signal elements as the full stimulus, see Table 9.1). For the signal only stimuli the predicted threshold elevation is a factor of $\sqrt{2}$, as there are no

interstitial noise elements integrated to limit performance in that condition. Models with this strategy capture the performance for the small check sizes (1 to 3) in the noise check condition, the medium check sizes (5 to 9) in the signal only condition and the largest black check size (15) in both the noise check and signal only conditions (see Figure 9.12a-b).

The "sum ideally" strategy (SI) involves combination of orientation information only from potential signal elements (see Figure 9.10). Models with this strategy predict threshold elevation of a factor of approximately $\sqrt{2}$ for both of the checked conditions (Figure 9.7), which is identical to the prediction for the signal only stimuli with the SA model. For this reason, the data in the signal only condition which were well-described by SA models are fit just as well by the SI models (see Figure 9.12c-d). In addition, the SI models predict the $\sqrt{2}$ summation for the medium check sizes in the noise check condition.

9.5.6 Two-stage hybrid models

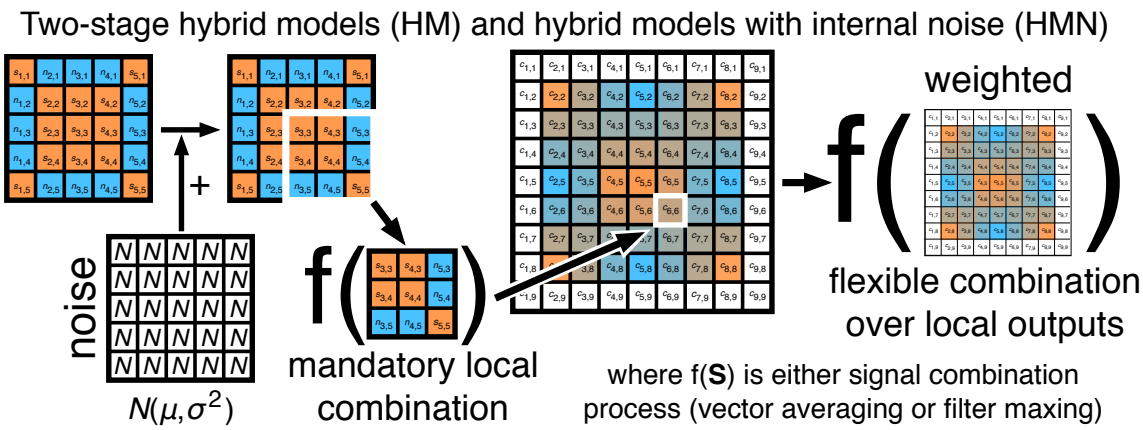

Figure 9.11: Showing the pooling strategies used in the two-stage "hybrid models" (HM) and hybrid models with internal noise (HMN). In the HMN models the representation of the stimulus is perturbed by independent Gaussian noise added at each location (including blank regions). In the HM models no noise is added. Following this, mandatory local signal combination takes place over 3 by 3 regions to create a new matrix where each element represents the pooled information from a local region in the original image (at this stage behaving according to the SA strategy). This is then followed by flexible combination of information over that matrix, with the contribution from each element weighted according to its expected signal to noise ratio (at this stage behaving according to the SI strategy).

The fact that the simpler candidate models featuring the two different pooling strategies (SA and SI) each predicted performance for different subsets of the data suggested that a complete account could be provided by a model that combined their two behaviours. In the two-stage "hybrid models" (HM; see Figure 9.11) mandatory local combination over a 3 by 3 region (behaving like the SA model) is followed by flexible pooling of the outputs from those regions weighted by their expected signal to noise ratios which were applied as a template (as in the

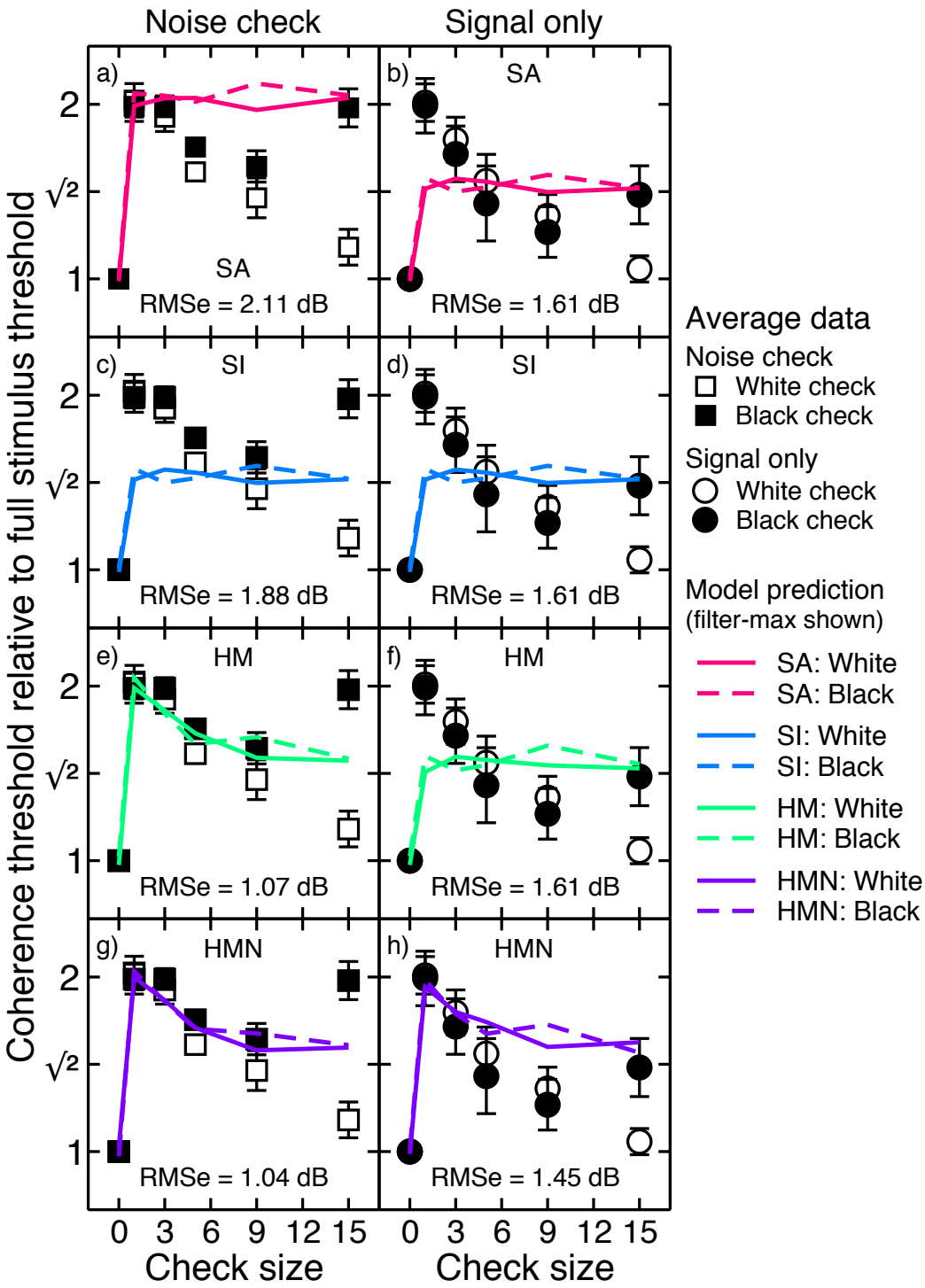

Figure 9.12: Coherence threshold data plotted with model predictions from the filter maxing versions of the SA (a-b), SI (c-d), HM (e-f) and HMN (g-h) models. Each row shows the same averaged dataset replotted from Figure 9.5. The only fitted parameter was the size of the pooling regions in the HM and HMN models (3 by 3), which was performed by hand. RMS errors between the model predictions and data are shown in dB.

SI model). The size of the local pooling region was the only parameter in this model. It was adjusted by hand to fit the data. In the noise check condition this model predicts an initial doubling of the threshold coherence for the small check sizes, followed by an improvement in performance to approach a factor of $\sqrt{2}$ for the medium and large check sizes (see Figure 9.7). This captures the performance for all but the largest (15) check sizes (Figure 9.12e-f). In the signal only condition the predictions are once again the same as for models with the SA and SI strategies, as there is still no additional noise to limit performance for the smaller check sizes.

9.5.7 Internal noise

The SA, SI and HM models all predict identical performance for the signal only condition. This is because there are no interstitial noise elements in that condition that can be inappropriately pooled to elevate the threshold coherence. The data however show a doubling of threshold for the small check sizes that is the same in the noise check and signal only conditions. This is a larger performance deficit than can be accounted for by any of the three models (which all predict a $\sqrt{2}$ threshold increase for this condition). A two-stage hybrid model featuring additive Gaussian internal noise at each location (HMN) does however predict similar performance for the signal only conditions as for the noise check conditions (see Figure 9.7). Because of the mandatory combination rule at the first pooling stage, this early noise model is equivalent to a model where performance is limited by noise affecting the output of the first pooling stage. This noise is discussed further in Section 9.6.2.

The standard deviation of the internal noise is now an additional parameter in the model, serving only to increase or decrease the sensitivity equally for all conditions. Calculating the relative thresholds from the output of the model therefore factors out the effects of this model parameter (meaning that it does not need to be adjusted to fit the model to the data). The predictions from the HMN model account for the average human thresholds in the noise check condition for all but the largest check size (Figure 9.12g-h). In the signal only condition thresholds are lower than that predicted by the model for the medium (9) and largest (15) white condition, but for all other conditions the human behaviour is captured by the model. The underperformance of the model compared to the data for the medium check size stimuli is driven entirely by the results from observers RJS and SAW (see Figure 9.5b), omitting these two observers from the analysis (not shown) raises the thresholds from these conditions into close agreement with the model.

9.5.8 Effects of eccentricity

In the HMN models, there are several possible reasons why thresholds could be elevated when signal is presented only in the periphery (compare black and white check thresholds for the largest check size stimulus in Figure 9.13). One possibility would be if the observer was only

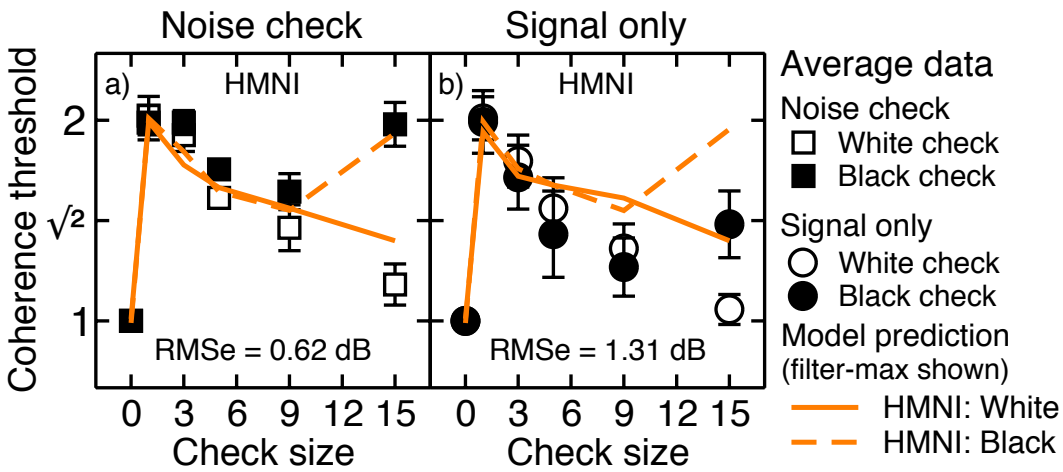

Figure 9.13: Coherence threshold data (replotted from Figure 9.5) plotted with model predictions from the filter maxing version of the HMNI model. The mandatory pooling region was set to 3 by 3 elements, and the aperture beyond which elements could not be pooled was set to 19 by 19 elements. These two parameters were fitted to the data by hand. RMS errors between the model predictions and data are shown in dB.

able to pool information from elements in the centre of the display. This was tested using a "hybrid model with internal noise *and inhomogeneity*" (HMNI), where the model observer only had access to information from elements that were within a central 19 by 19 element square aperture (equivalent to 19 degrees or 76 carrier cycles). This aperture size provided the best fit to the data (compared to other square apertures with integer dimensions). Figure 9.13 shows the prediction from this model plotted against the data from the noise check and signal only conditions. The HMNI model prediction is similar to that made by the HMN model for all conditions except for the largest check size (15). The data from the noise check condition are well-described by this model, providing a better fit than the HMN model (0.62 dB vs. 1.04 dB). The fit is also superior for the signal only condition (1.31 dB vs. 1.45 dB), however the model predicts that thresholds should be higher in the largest (15) black check condition than they are in the data (predicting the same threshold for the noise check and signal only condition). Alternative explanations for the divergence in performance for the largest check size are considered in the Discussion (Section 9.6.3).

9.6 Discussion

9.6.1 Orientation integration is a noisy two-stage process

The results of this study suggest that the combination of orientation information over space is a noisy two-stage process (Figure 9.14). The results here can be accounted for by a model that performs mandatory local integration affected by internal noise at each location (this noise is discussed further below) followed by flexible pooling over the outputs from those regions.

This is the case regardless of the assumed signal combination process for the two considered here (vector averaging versus maxing over filter outputs). These results are in agreement with previous studies that have found lower thresholds for stimuli with a greater signal area (Dakin, 2001; Jones et al., 2003). The effect found for the arrangement of the elements in the display (which forms the basis for the proposed noisy two-stage model) contradicts the flexibility attributed to the pooling of local samples Dakin (2001). Concerning the information limit theory presented in that study, it is not possible to determine whether the observers here are pooling across the entire stimulus or making their decision based on only a small proportion of the local samples in the display (in accordance with the information limit theory).

It is not entirely clear how the results presented here can be reconciled with those of Husk et al. (2012), who found no summation with increasing signal area for similar stimuli. The main difference between the stimuli used in the two studies is that this study used the Battenberg summation paradigm, whereas Husk et al. (2012) increased the signal area of their stimuli by increasing their diameter. It is possible that a combination of decreasing sensitivity for the local orientation discrimination task and increasing mandatory summation region size with eccentricity might flatten the threshold versus area functions (see below).

9.6.2 The nature of the limiting internal noise

The similarity of the results from the noise check and signal only conditions suggests that the noise limiting performance in this study does not occur in the response to individual elements. Instead, these results suggest that a level of noise is pooled that is proportional to the number of locations being monitored, including blank locations which are being mandatorily integrated at the first stage of the two-stage model. Jones et al. (2003) included similiar "late noise" when modelling data from an orientation coherence experiment that used filtered noise as stimuli, however in the modelling for that study the late noise was constant for different signal areas (this would be N_{late} in Figure 9.14). Such a model would not explain the results from this study, as the improvement in performance seen for the medium check size stimuli (in both the noise check and signal only conditions) requires that the observer is able to segregate out the limiting noise in irrelevant regions from the second combination stage (if dominant noise is contributed from each monitored location then this could be performed by weighting the local outputs according to a template w, as shown in Figure 9.14).

In the modelling performed here (see Figure 9.14) the noise was implemented at each pooled location after the initial filtering stage (N_{early}), however due to the mandatory local combination at the first combination stage this is equivalent to adding noise to the combined local outputs (N_{mid}). From the results of this study it is not possible to determine whether the limiting noise should be N_{early} or N_{mid} in Figure 9.14, however N_{mid} seems more plausible as the level of early noise needed to match or exceed the noise introduced by the randomly oriented elements in the noise check stimuli would be very high. There is also the possibility that the

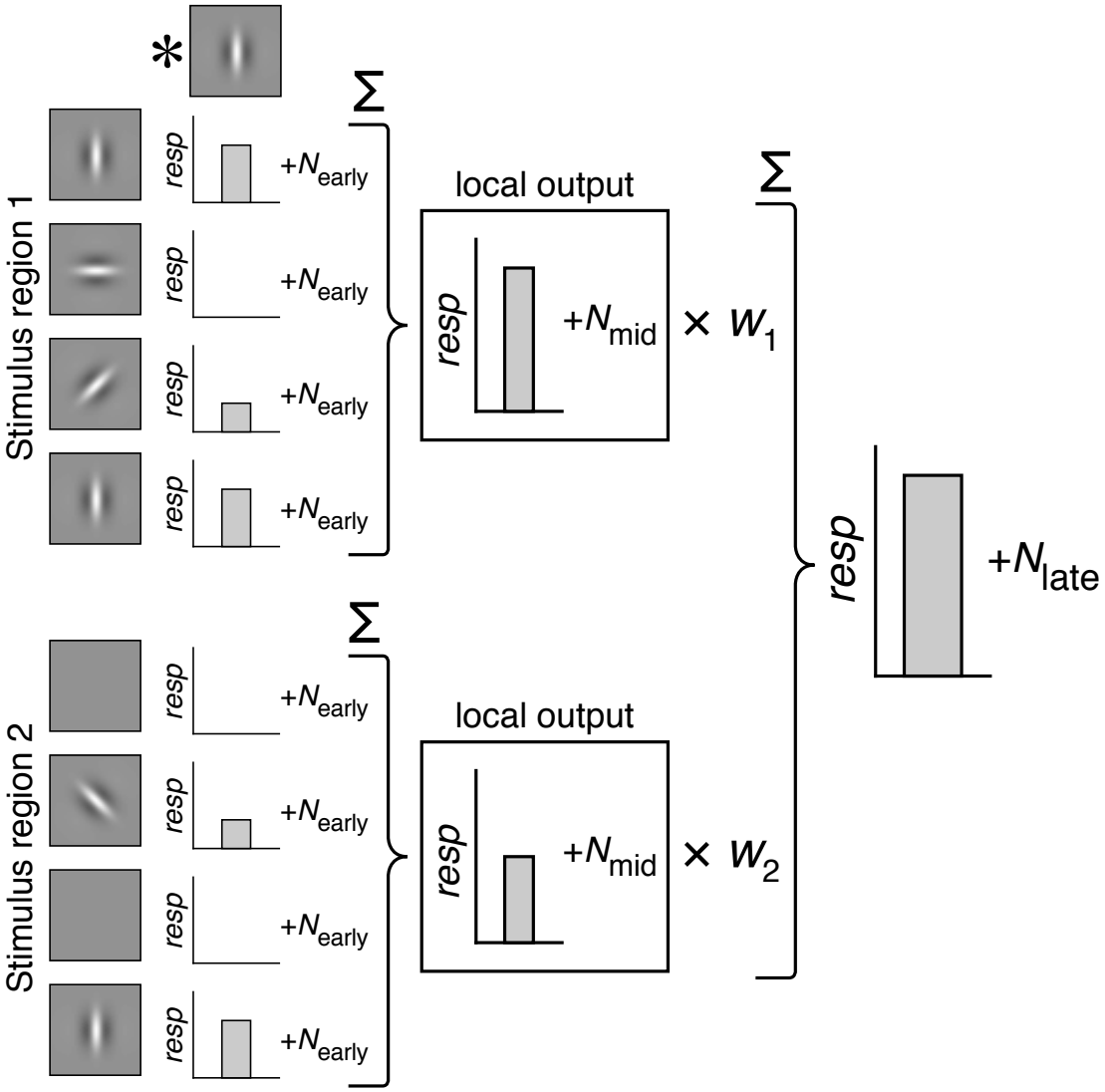

Figure 9.14: Diagram of the two-stage hybrid model with internal noise, using the filter maxing combination process. This diagram shows how the "vertical" response is determined by filtering with a vertical filter element, mandatory local summation, and then global summation of the local outputs weighted by the expected signal-to-noise ratio (w). The "horizontal" response would be calculated in an identical manner, except with a horizontal filter element at the convolution stage. N_{early}, N_{mid}, and N_{late} show three possible locations for the limiting internal noise (discussed in the text).

limiting internal noise could be multiplicative, rather than the additive noise implemented in the modelling here. Modifying the stimuli used in this study to perform an equivalent noise experiment would allow for the properties of this noise to be investigated further.

Another open question is whether this noise is affecting filter responses within an oriented channel, which would impair performance when maxing across filter outputs (as in Figure 9.14), or if the noise is instead affecting some estimate of local orientation (as it would in a model that was performing a vector average to combine local samples). This chapter has compared the predictions from filter maxing and vector averaging models, where these combination pro-

cesses were used in *both* pooling stages (the local mandatory integration and the "global" flexible pooling of these local outputs) however the fact that there appear to be two stages brings up the possibility that one combination process may be used locally and another used globally. In future work the orientation Battenberg stimuli used in this study could be adapted to test this hypothesis and determine which combination process is used at each of the two stages.

9.6.3 Effects of eccentricity

The Battenberg stimuli used in this study were designed with the intention of factoring out the effects of eccentricity. This was largely successful as performance was similar for the black and white check phases for all stimuli except for the largest check size. In this case performance diverged, with the coherence thresholds for the white check stimuli decreasing to around the level for the full stimulus (meaning that approximately half as many signal elements in the entire display were required for detection, see Figure 9.6a-b). In the noise check condition the threshold for the black checks was approximately twice as high as for the full stimulus, indicating that the same number of signal elements in the entire display were required for detection (Figure 9.6a). In the signal only condition the performance deficit for the black checks condition was much smaller.

One explanation for these data would be if local sensitivity for performing the orientation task was poorer in the periphery (Jones et al., 2003). Restricting the integration region to the central 19 by 19 elements provided a good fit to the noise check data, but predicted too much of a performance deficit for the largest black check stimulus in the signal only condition. In fact, any model that accounts for the noise check data using a sensitivity decline with eccentricity would make the same prediction for the signal only condition. The difference in performance between the noise check and signal only conditions for the largest black check stimulus suggests that if performance is limited by an inhomogeneity in sensitivity then the observer must be exploiting some other process to improve performance in the signal only condition. One possible way in which this could be done would be if there were multiple apertures available after the second combination stage that the observer then combines information over, making a *three-stage* hybrid model. In order to make a distinct prediction from the two-stage hybrid model, this third combination process would have to follow another nonlinearity (one possibility would be probability summation over multiple apertures, after Baker & Meese, 2011).

Alternatively, the maximum integration aperture at the second stage could be a limited number of elements rather than a spatial limit (although this would predict different thresholds in the noise check and signal only conditions for the smallest check sizes, an effect that is not seen in the data). Other possiblities would include an increase in the size of the mandatory local integration region with eccentricity (making this local combination stage equivalent to crowding, see Parkes et al., 2001), or constraints on the flexible pooling that prevent an ideal

strategy being adopted for this condition (for example, pooling may not occur around an annulus). Although these questions are important for the development of a full account of how orientation signals are integrated over space, this study is not designed in such a way as to constrain any analysis of the effects of eccentricity on the pooling (instead being designed to factor out these effects where possible).

CHAPTER 10

Discussion

> "To know the brain... is equivalent to ascertaining the material course of thought and will, to discovering the intimate history of life in its perpetual duel with external forces."
> – *Santiago Ramón y Cajal*

THE work presented in the five experimental chapters of this thesis provides an account of how signals are combined over space in the human visual system, including the effects of the inhomogeneous sensitivity to contrast across the visual field. The summation of contrast over space in amblyopia is also addressed, as well as the summation of coherent orientation signals. In this chapter I summarise my findings and suggest future studies that I aim to carry out in order to further develop this work.

10.1 Conclusions from the work presented here

10.1.1 The visual field inhomogeneity in log contrast sensitivity is bilinear

Previously, the decline in log contrast sensitivity for grating patches was understood to be a linear function of eccentricity. For a wide range of spatial frequencies (1.6 - 12.8 c/deg) the slope of the decline was constant when eccentricity was expressed in stimulus carrier cycles (Pointer & Hess, 1989). The study presented in Chapter 4 establishes that the decline in log contrast sensitivity in the central visual field (within 18 cycles of fixation) is better characterised as bilinear, initially steep before then shallowing to around half that gradient. As a 3D surface that is a function of both horizontal and vertical eccentricity, this bilinear sensitivity decline takes the form of a witch's hat (Baldwin et al., 2012).

Like the Pointer and Hess (1989) result, the bilinear sensitivity functions reported in this thesis are scale invariant across a wide range of spatial frequencies (0.7 - 4.0 c/deg) and show a shallower decline across the horizontal meridian than across the vertical meridian. The superior and inferior declines of the vertical meridian were asymmetric, with the second slope of the inferior decline being more shallow than that of the superior decline. This is in agreement with the vertical meridian asymmetry reported previously (for a review, see Abrams et al., 2012). The study conducted here did not find any strong or consistent effect of stimulus orientation within the central visual field, indicating that these results can be generalised across different orientations.

The attenuation surface developed in Chapter 4 provides a more accurate account of the contrast sensitivity decline within the central visual field than has been reported previously. It is used elsewhere in this thesis both in models of the visual system (Chapter 5) and to transform stimuli in order to counteract the effects of the inhomogeneity in sensitivity (see below). It is not clear what physiological feature or features of the visual system might cause the scale invariant bilinear decline reported here. Ahumada and Watson (2011) suggested that sensitivity could be predicted from the retinal cone density functions (see Section 4.7.6). Although this correctly predicts the concave shape of the log contrast sensitivity declines, a model based on cone density would fail to predict the scale invariance found here. In order for the effect to be scale invariant, the strongest part of the inhomogeneity must occur where neurones are spatial frequency selective. The earliest point at which this is seen is the ganglion cell layer in the retina. Recent work by Bradley and Geisler (2012) has attempted to build a bottom-up model of visual processing starting from the ganglion cell stage. I have been in contact with the authors to try to establish whether this model can predict the witch's hat attenuation surface.

10.1.2 Area summation is spatially extensive and occurs according to a single rule

By applying the inverse of the attenuation surface to the stimuli Chapter 6 demonstrates that the summation of threshold contrast occurs under a single (fourth-root) rule over at least 33 cycles once the visual field inhomogeneity in contrast sensitivity is factored out. These data were fit by a variety of summation models (developed in Chapter 5). The summation slopes found were consistent with either a probability summation model or noisy energy model. The results of previous studies (e.g. Meese & Summers, 2012) lead me to favour the noisy energy model interpretation of the results. As the final stage of the noisy energy model is a linear sum performed over the local filtered, transduced, and template-weighted responses to the stimulus image, this suggests that either i) a linear summing mechanism exists in the visual system that combines the outputs from simple cells over at least 33 cycles, or ii) several such mechanisms of smaller sizes (e.g. complex cells) exist, whose outputs are then also combined by a linear sum over at least 33 cycles (a series of two linear summation operations being indistinguishable from a single linear sum).

The findings in Chapter 7 show that area summation occurs according to a single (fourth-root) summation rule in the fovea, parafovea, and periphery. As above, the noisy energy model is favoured over the probability summation model based on the results from other studies (e.g. Meese & Summers, 2012). This finding suggests that the most of the differences in the results from summation experiments conducted at different locations in the visual field would be accounted for by the visual field inhomogeneity in contrast sensitivity. After the initial attenuation and filtering stages, the responses from the detectors appear to have been combined by a common set of processes at the three different locations tested in this study. One observer did show some additional summation in the parafovea which may have been caused by elongation of the filter elements (simple cell receptive field analogues) with eccentricity, and another observer showed some unusually high thresholds for very small stimuli in the periphery, however in general the model provided an excellent account of the data with a single fitted parameter (global sensitivity) for each observer.

10.1.3 Summation of threshold contrast over area is normal in amblyopia

Spatial summation in amblyopia had previously been explained by a probability summation model (Hess & Howell, 1978; Hess & Campbell, 1980), however the rejection of this model as an account of summation in the normal visual system (in favour of a noisy energy model) prompted me to investigate whether summation in amblyopia also needs to be reevaluated. The "Battenberg" study on area summation in amblyopia presented in Chapter 8 shows that spatial summation is normal in amblyopia in at least some cases and is best explained by a model which features linear filtering followed by a linear sum of the squared filter outputs. This is consistent with the noisy energy model if it cannot match a template to the contrast modulation in the Battenberg stimulus and instead pools contrast over the stimulus extent (as was reported in the original study conducted in normal observers by Meese, 2010).

It is suspected that the results for some of the observers in this study were confounded by artefacts introduced by the checkerboard modulation of the Battenberg stimuli, leading to anomalous short-range summation results for the normal observer at medium spatial frequencies (4 - 8 c/deg) and for some of the amblyope observers. This effect could only reduce the amount of measured summation however, and so could not account for the cases where greater summation was found than that predicted by the probability summation model (as is the case for half of the amblyopes tested in Chapter 8). From this finding it appears that amblyopes show the same spatial summation behaviour as normal observers. This suggests that the higher-level mechanisms that sum the responses from detectors positioned across the visual field are intact in amblyopia.

10.1.4 The summation of orientation signals is a noisy two-stage process

By extending the Battenberg stimulus paradigm to the orientation domain (Chapter 9), I have found that the spatial configuration of the signal areas in a stimulus has a significant effect on its coherence threshold. This is surprising because previous studies have indicated that the arrangement of the stimulus elements does not affect performance (Dakin, 2001). For the stimuli that contained a checkerboard pattern of potential signal and noise regions ("noise check" condition), thresholds initially doubled compared to the "full" stimulus for the smallest check sizes (consistent with the observer combining information from all elements in the stimulus) before decreasing to a factor of approximately $\sqrt{2}$ above the full stimulus threshold for the medium check sizes (consistent with the observer only monitoring potential signal elements).

The data from the noise check experiment were well-described by a two-stage "hybrid model" that first performed a mandatory local integration of elements within a fixed radius before then combining the outputs of that first stage ideally. The only data points not fit by this model were those from the largest check size, where performance diverged dependent on whether most of the signal was presented to the fovea or in the periphery. Limiting the observer to only use information from the elements presented within a square 19 by 19 degree (76 by 76 stimulus carrier cycle) aperture in the centre of the display provided a much better fit to the data. This is similar to the result from a previous experiment on the summation of threshold contrast over large areas that found a maximum integration region (Baker & Meese, 2011). Note that the aperture size proposed here is much larger than the region tested in the area summation experiment presented in Chapter 6, which found no limit on the summation of contrast over 33 stimulus carrier cycles.

Coherence thresholds for Battenberg stimuli were similar when the noise regions were replaced by blank space ("signal only" condition). This was an unexpected finding, as based on the information from the stimulus alone a $\sqrt{2}$ threshold elevation would be expected for all of the Battenberg conditions (as there are no noise elements in the non-signal regions to impair performance). This indicates that the noise introduced by the randomly oriented elements in those regions was not the limiting factor on performance. This result would be consistent with performance for the task being limited by internal noise that is proportional to the area being monitored, regardless of the element density in that area. This can be built into the hybrid two-stage model by adding internal noise at each location in the input stage, however the mandatory summation in the first stage of the model means that the limiting internal noise could also be placed *between* the first and second stages. Further work is needed to explore the nature of the limiting internal noise, and to determine whether this model can be adjusted to account for the differences in performance for the noise check and signal only conditions when signal is presented to the periphery.

10.2 Future work

10.2.1 Summation of contrast over space

Together with other recent work (Meese & Summers, 2007; Meese, 2010; Baker & Meese, 2011; Meese & Summers, 2012), the results reported in this study provide strong support for a noisy energy model interpretation of area summation across the visual field. There are still some questions left unanswered however. The first of these is whether there is in fact a limit on the extent over which contrast can be combined. Chapter 6 finds that two out of three observers were able to combine signal over at least 33 cycles, whereas the third observer's performance asymptoted at 19 cycles. It is not clear whether this observer might be capable of summing over a greater area under different experimental conditions. For the other two observers, there is an open question of whether there is maximum extent over which summation can occur. An area summation experiment conducted with witch hat compensated gratings having a much wider range of stimulus sizes could reveal whether such a limit exists.

Using larger stimuli may also shed some light on whether the more accurate form of the noisy energy model includes a weighting of the template to account for the visual field inhomogeneity in contrast sensitivity. The studies presented in Chapters 6 and 7 show a slight preference for a noisy energy model with a "flat" template (i.e. without any weighting by the attenuation surface) however the design of these studies causes the predictions from the two models to be very difficult to distinguish. For larger non-compensated gratings the inefficiency of monitoring the peripheral regions of the stimulus would be sufficient to separate the predictions made by models featuring the two different kinds of template.

10.2.2 Integration of orientation signals

The orientation integration study conducted here raises several questions that could be addressed in future studies. The first of these is whether the method used to perform the integration over the Battenberg stimuli is maxing over filter outputs or whether it is performing another operation such as vector averaging. In particular, it would be interesting to see whether behaviour might switch from one strategy to the other between the "full" and "checked" stimuli. Although these different integration processes made the same prediction for the study presented in this thesis, they could be distinguished between by investigating performance when the orientation of the noise elements is biased (Webb et al., 2007).

Although the study presented can draw some conclusions about how elements are pooled over space to calculate a global orientation, a more complete account of this process could be uncovered through the use of psychophysical reverse correlation (Neri, 2002). This method involves keeping track of stimulus information for each trial along with a record of the observer's

response in order to work out which features of the stimulus the observer was responding to. Future work could also investigate the nature of the noise that limits performance for this task. The study presented in Chapter 9 suggests that performance is not limited by noise affecting the response to each element, but instead to noise that varies with the monitored signal area (whether it contains elements or not). An equivalent noise method (similar to that used by Dakin, 2001) could be used to determine the nature of the limiting noise for this task.

10.2.3 Extending the orientation Battenberg work to the motion domain

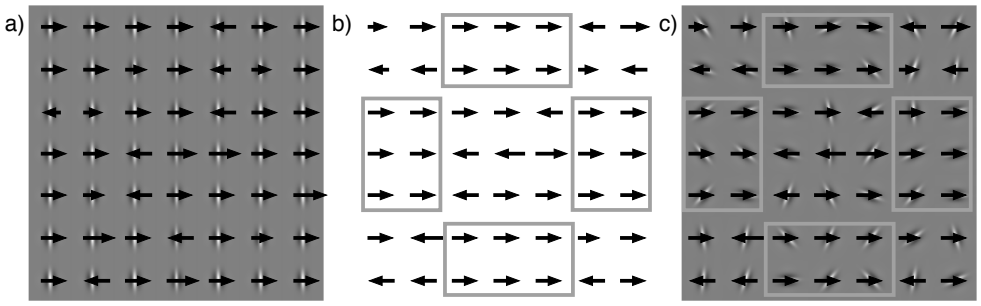

Figure 10.1: Preliminary design for motion Battenberg stimuli (a). The phase of each element will rotate to cause the perception of local motion in that element (motion vectors are shown by the black arrows). This can then give coherent motion in elements that are in the signal regions of the checkerboard pattern (b, grey boxes indicate signal regions, background image omitted for clarity). This can also be done for stimuli where elements have different orientations to give a complex motion task (c). For this stimulus the rate of phase rotation in signal elements will depend on their orientation, to give the impression of a single patterned object moving in one direction behind a grid of apertures.

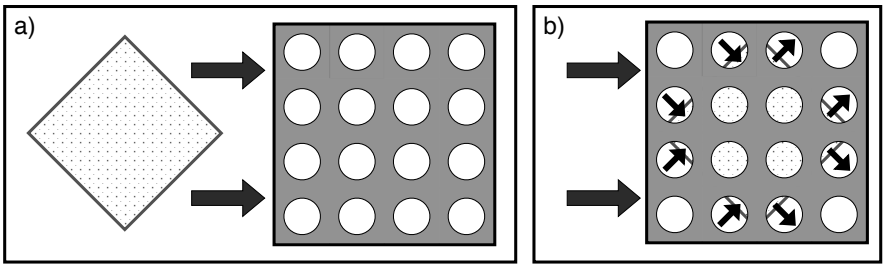

Figure 10.2: Demonstrating the aperture problem for identifying the motion direction of an object from its local motion vectors. The object in panel a) moves behind a screen with many apertures in panel b). The movement of the object's edges seen within each aperture do not necessarily correspond to the direction of the motion of the object. For any single aperture these local motion vectors could correspond to several possible global motion directions (this is the aperture problem). Multiple motion vectors must be combined over space in order to calculate the true trajectory of the object. As each V1 neurone only receives information from a small part of the visual field these calculations must be performed by the visual system to determine the trajectory of any object.

The orientation Battenberg experiment can be considered to test global form perception. An

analogous series of experiments could be designed to test global motion. The detection of global motion is thought to occur in a separate extra-striate region of the brain to that of global form (Mishkin & Ungerleider, 1982; Goodale & Milner, 1992; Braddick et al., 2000), however the pooling involved in each may result from the outputs of common mechanisms or involve similar strategies. A comparison could be made between the global processing of form and that of motion by designing Battenberg stimuli that differ only in whether they contain coherent form or coherent motion. These stimuli could be used to investigate the extra-striate deficit reported in amblyopia (Simmers et al., 2003; Simmers, Ledgeway, & Hess, 2005; Hess, Mansouri, Dakin, & Allen, 2006).

Additional studies could then be performed on the integration of complex signals (formed from more than one component signal) over area. In the motion domain this processing is required to overcome the aperture problem, where the local direction of motion in a stimulus may correspond to several possible directions of global motion for the complete pattern that it is sampled from. Some spatial properties of this combination for motion stimuli have been previously examined in neurophysiology (Majaj, Carandini, & Movshon, 2007) and psychophysics (Amano, Edwards, Badcock, & Nishida, 2009), however the spatial limits of the integration of these complex signals have not yet been addressed. A Battenberg stimulus could be used to investigate the spatial properties of the signal combination.

10.3 Conclusion

The findings reported in this thesis add several key features to the understanding of how the visual system combines signals over area. The witch's hat attenuation surface provides a more accurate model of how contrast sensitivity varies as a function of eccentricity than has been reported previously. The spatial summation studies demonstrate for the first time that stimulus contrast is summed over at least 33 grating carrier cycles according to a fourth-root summation rule (interpreted as a noisy energy model here) and that this same summation rule is followed in the fovea, parafovea, and periphery. The Battenberg summation study conducted on amblyopes suggests that this same noisy energy model describes how contrast is combined over the visual field in the amblyopic visual system. Finally, extending the methods used for the contrast studies here to the combination of orientation signals across the visual field has revealed that the configuration of signals in the stimulus has an effect on the detectability of coherent orientation. I have developed a noisy two-stage combination model in order to explain the human performance for this task.

References

Abrams, J., Nizam, A., Carrasco, M., 2012. Isoeccentric locations are not equivalent: The extent of the vertical meridian asymmetry. Vision Research 52 (1), 70–78.

Ahumada, A. J., Watson, A. B., 2011. Updating the Standard Spatial Observer for contrast detection. Investigative Opthalmology & Visual Science 52 (E-Abstract 1879).

Akaike, H., 1974. A New Look at the Statistical Model Identification. IEEE Transactions On Automatic Control 19 (6), 716–723.

Amano, K., Edwards, M., Badcock, D. R., Nishida, S., 2009. Adaptive pooling of visual motion signals by the human visual system revealed with a novel multi-element stimulus. Journal of Vision 9(3) (4), 1–25.

Anderson, S. J., Holliday, I. E., Harding, G. F. A., 1999. Assessment of cortical dysfunction in human strabismic amblyopia using magnetoencephalography (MEG). Vision Research 39, 1723–1738.

Anderson, S. J., Mullen, K. T., Hess, R. F., 1991. Human peripheral spatial resolution for achromatic and chromatic stimuli: limits imposed by optical and retinal factors. Journal of Physiology 442, 47–64.

Baker, D. H., Meese, T. S., 2011. Contrast integration over area is extensive: A three-stage model of spatial summation. Journal of Vision 11:14 (14), 1–16.

Baker, D. H., Meese, T. S., 2012. Using psychophysical reverse correlation to measure the extent of spatial pooling of luminance contrast. Perception 41 (1512).

Baldwin, A. S., Meese, T. S., 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.

Barlow, H. B., 1958. Temporal and spatial summation in human vision at different background intensities. Journal of Physiology 141, 337–350.

Barnes, G. R., Hess, R. F., Dumoulin, S. O., Achtman, R. L., Pike, G. B., 2001. The cortical deficit in humans with strabismic amblyopia. Journal of Physiology 533 (1), 281–297.

Berkley, M. A., Kitterle, F., Watkins, D. W., 1975. Grating visibility as a function of orientation and retinal eccentricity. Vision Research 15 (2), 239–244.

Bijl, P., Koenderink, J. J., Koenderink, P. B. J. J., 1993. Visibility of Elliptical Gaussian Blobs. Vision Research 33 (2), 243–255.

Blakemore, C., Campbell, F. W., 1969. On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. Journal of Physiology 203, 237–260.

Blakemore, C., van Sluyters, R. C., 1974. Reversal of the physiological effects of monocular deprivation in kittens: further evidence for a sensitive period. Journal of Physiology 237, 195–216.

Blakemore, C., Vital-Durand, F., 1986. Effects of visual deprivation on the development of the monkey's lateral geniculate nucleus. Journal of Physiology 380, 493–511.

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

Braddick, O. J., O'Brien, J. M. D., Wattam-Bell, J., Atkinson, J., Turner, R., 2000. Form and motion coherence activate independent, but not dorsal/ventral segregated, networks in the human brain. Current Biology 10 (12), 731–734.

Bradley, C., Geisler, W. S., 2012. A model of target detectability across the visual field in naturalistic backgrounds. Journal of Vision 12 (9), 318–318.

Brainard, D. H., 1997. The Psychophysics Toolbox. Spatial Vision 10 (4), 433–436.

Bullier, J., Norton, T. T., 1979. Comparison of Receptive-Field Properties of X and Y Ganglion Cells With X and Y Lateral Geniculate Cells in the Cat. Journal of Neurophysiology 42 (1), 274–291.

Campbell, F. W., Kulikowski, J. J., 1966. Orientational selectivity of the human visual system. Journal of Physiology 187, 437–445.

Campbell, F. W., Kulikowski, J. J., Levinson, J., 1966. The effect of orientation on the visual resolution of gratings. Journal of Physiology 187, 427–436.

Campbell, F. W., Robson, J. G., 1968. Application of Fourier analysis to the visibility of gratings. Journal of Physiology 197, 551–566.

Carandini, M., Heeger, D. J., Movshon, J. A., 1999. Linearity and gain control in V1 simple cells. In: Ulinski, P. S., Jones, E. G., Peters, A. (Eds.), Cerebral cortex, Vol 13, Models of cortical circuits. Kluwer Academic/Plenum, New York, pp. 401–443.

Corso, J. F., 1963. A theoretico-historical review of the threshold concept. Psychological Bulletin 60 (4), 356–370.

Curcio, C. A., Allen, K. a., 1990. Topography of ganglion cells in human retina. The Journal of Comparative Neurology 300 (1), 5–25.

Curcio, C. A., Sloan, K. R., Kalina, R. E., Hendrickson, A. E., 1990. Human photoreceptor topography. The Journal of Comparative Neurology 292, 497–523.

Dakin, S. C., 2001. Information limit on the spatial integration of local orientation signals. Journal of the Optical Society of America A 18 (5), 1016–1026.

Dakin, S. C., Watt, R. J., 1997. The Computation of Orientation Statistics from Visual Texture. Vision Research 37 (22), 3181–3192.

Daniel, P. M., Whitteridge, D., 1961. The representation of the visual field on the cerebral cortex in monkeys. Journal of Physiology 159, 203–221.

Davis, E. T., Graham, N., 1981. Spatial frequency uncertainty effects in the detection of sinusoidal gratings. Vision Research 21, 705–712.

Davis, E. T., Kramer, P., Graham, N., 1983. Uncertainty about spatial frequency, spatial position, or contrast of visual patterns. Perception & Psychophysics 33 (1), 20–28.

De Valois, K. K., De Valois, R. L., Yund, E. W., 1979. Response of striate cortex cells to grating and checkerboard patterns. Journal of Physiology 291, 483–505.

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

De Valois, R. L., De Valois, K. K., 1990a. Retinogeniculate Anatomy and Physiology. In: Spatial Vision. Oxford University Press, Oxford, Ch. 3, pp. 55–93.

De Valois, R. L., De Valois, K. K., 1990b. Spatial Vision. Oxford University Press, Oxford.

De Valois, R. L., De Valois, K. K., 1990c. Striate Cortex. In: Spatial Vision. Oxford University Press, Oxford, Ch. 4, pp. 94–146.

Dixon, W. J., Mood, A. M., 1948. A Method for Obtaining and Analyzing Sensitivity Data. Journal of the American Statistical Association 43 (241), 109–126.

Downing, C. J., Movshon, J. A., 1989. Spatial and temporal summation in the detection of motion in stochastic random dot displays. Supplement to Investigative Opthalmology and Visual Science 30, 72–87.

Finney, D. J., 1971. Probit analysis. Cambridge University Press, Cambridge, UK.

Flynn, J. T., 1967. Spatial Summation in Amblyopia. Archives of Ophthalmology 78 (4), 470–474.

Foley, J. M., 1994. Human luminance pattern-vision mechanisms: masking experiments require a new model. Journal of the Optical Society of America A 11 (6), 1710–1719.

Foley, J. M., Legge, G. E., 1981. Contrast detection and near-threshold discrimination in human vision. Vision Research 21, 1041–1053.

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

Gabor, D., 1946. Theory of communication. Journal of the Institution of Electrical Engineers - Part III: Radio and Communication Engineering 93 (26), 429–441.

García-Pérez, M. A., 1998. Forced-choice staircases with fixed step sizes: asymptotic and small-sample properties. Vision Research 38, 1861–1881.

García-Pérez, M. a., Peli, E., 2001. Luminance artifacts of cathode-ray tube displays for vision research. Spatial Vision 14 (2), 201–215.

García-Pérez, M. A., Sierra-Vázquez, V., 1996. Do channels shift their tuning towards lower spatial frequencies in the periphery? Vision Research 36 (20), 3339–3372.

Goodale, M. A., Milner, A. D., 1992. Separate visual pathways for perception and action. Trends in Neurosciences 15 (1), 20–25.

Graham, C. H., Brown, R. H., Mote, F. A., 1939. The relation of size of stimulus and intensity in the human eye: I. Intensity thresholds for white light. Journal of Experimental Psychology 24 (6), 555–573.

Graham, N., 1981. The visual system does a crude fourier-analysis of patterns. SIAM-AMS Proceedings 13, 1–16.

Graham, N., Robson, J. G., Nachmias, J., 1978. Grating summation in the fovea and periphery. Vision Research 18, 815–825.

Graham, N. V. S., 1989a. Extrinsic Uncertainty and Summation Revisited. In: Visual Pattern Analyzers. Oxford University Press, Oxford, Ch. 7, pp. 249–285.

Graham, N. V. S., 1989b. Intrinsic Uncertainty and Transducer Functions. In: Visual Pattern Analyzers. Oxford University Press, Oxford, Ch. 8, pp. 286–312.

Graham, N. V. S., 1989c. Neurophysiology and Psychophysics. In: Visual Pattern Analyzers. Oxford University Press, Oxford, Ch. 1, pp. 3–34.

Graham, N. V. S., 1989d. Some Mathematics. In: Visual Pattern Analyzers. Oxford University Press, Oxford, Ch. 2, pp. 35–86.

Green, D. M., 1961. Detection of Auditory Sinusoids of Uncertain Frequency. The Journal of the Acoustical Society of America 33 (7), 897–903.

Green, D. M., Swets, J. A., 1966a. Assumed Distribution of Signal and Noise. In: Signal Detection Theory and Psychophysics. John Wiley & Sons, New York, Ch. 3, pp. 53–85.

Green, D. M., Swets, J. A., 1966b. Elements of Statistical Decision Theory. In: Signal Detection Theory and Psychophysics. John Wiley & Sons, New York, Ch. 1, pp. 7–29.

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

Green, D. M., Swets, J. A., 1966d. Statistical Decision Theory and Psychophysical Procedures. In: Signal Detection Theory and Psychophysics. John Wiley & Sons, New York, Ch. 2, pp. 30–52.

Green, D. M., Swets, J. A., 1966e. The Sensory Threshold and Psychophysical Method. In: Signal Detection Theory and Psychophysics. John Wiley & Sons, New York, Ch. 5, pp. 117–148.

Gregory, R., 1966. Eye and Brain: The Psychology of Seeing. Weidenfeld and Nicolson, London.

Heeger, D. J., 1991. Nonlinear Model of Neural Responses in Cat Visual Cortex. In: Landy, M., Movshon, J. A. (Eds.), Computational Models of Visual Processing. MIT Press, Cambridge, MA, pp. 119–133.

Heeger, D. J., 1992. Normalization of cell responses in cat striate cortex. Visual Neuroscience 9, 181–197.

Heeley, D., Timney, B., 1988. Meridional anisotropies of orientation discrimination for sine wave gratings. Vision Research 28 (2), 337–344.

Hess, R. F., 1980. A preliminary investigation of neural function and dysfunction in amblyopia—I. Vision Research 20 (9), 749–754.

Hess, R. F., Baker, D. H., May, K. A., Wang, J., 2008. On the decline of 1st and 2nd order sensitivity with eccentricity. Journal of Vision 8(1) (19), 1–12.

Hess, R. F., Campbell, F., 1980. A preliminary investigation of neural function and dysfunction in amblyopia—II. Vision Research 20 (9), 755–756.

Hess, R. F., Campbell, F. W., Greenhalgh, T., 1978. On the Nature of the Neural Abnormality in Human Amblyopia: Neural Aberrations and Neural Sensitivity Loss. European Journal of Physiology 377, 201–207.

Hess, R. F., Howell, E. R., 1977. The threshold contrast sensitivity function in strabismic amblyopia: evidence for a two type classification. Vision Research 17 (9), 1049–1055.

Hess, R. F., Howell, E. R., 1978. The influence of field size for a periodic stimulus in strabismic amblyopia. Vision Research 18, 501–503.

Hess, R. F., Mansouri, B., Dakin, S. C., Allen, H. A., 2006. Integration of local motion is normal in amblyopia. Journal of the Optical Society of America A 23 (5), 986–992.

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

Hoekstra, J., van der Groot, D. P. J., van den Brink, G., Bilsen, F. A., 1974. The influence of the number of cycles upon the visual contrast threshold for spatial sine wave patterns. Vision Research 14, 365–368.

Howell, E. R., Hess, R. F., 1978. The functional area for summation to threshold for sinusoidal gratings. Vision Research 18, 369–374.

Hubel, D. H., Wiesel, T. N., 1959. Receptive fields of single neurones in the cat's striate cortex. Journal of Physiology 148, 574–591.

Hubel, D. H., Wiesel, T. N., 1962. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. Journal of Physiology 160, 106–154.

Hubel, D. H., Wiesel, T. N., 1970. The period of susceptibility to the physiological effects of unilateral eye closure in kittens. Journal of Physiology 206, 419–436.

Husk, J. S., Huang, P.-C., Hess, R. F., 2012. Orientation coherence sensitivity. Journal of Vision 12:6 (18), 1–15.

Johnston, A., 1987. Spatial scaling of central and peripheral contrast-sensitivity functions. Journal of the Optical Society of America A 4 (8), 1583 – 1593.

Jones, D. G., Anderson, N. D., Murphy, K. M., 2003. Orientation discrimination in visual noise using global and local stimuli. Vision Research 43, 1223–1233.

Jones, J. P., Palmer, L. A., 1987. An Evaluation of the Two-Dimensional Gabor Filter Model of Simple Receptive Fields in Cat Striate Cortex. Journal of Neurophysiology 58 (6), 1233–1258.

Kaernbach, C., 2001. Slope bias of psychometric functions derived from adaptive data. Perception & Psychophysics 63 (8), 1389–1398.

Kelly, D. H., 1984. Retinal inhomogeneity. I. Spatiotemporal contrast sensitivity. Journal of the Optical Society of America A 1 (1), 107–113.

Kersten, D., 1984. Spatial summation in visual noise. Vision Research 24 (12), 1977–1990.

Kingdom, F. A. A., Prins, N., 2010. Psychophysics: A Practical Introduction. Academic Press, London.

Kleiner, M., Brainard, D. H., Pelli, D. G., 2007. What's new in Psychtoolbox-3? Perception 36 (ECVP Abstract Supplement).

Koenderink, J. J., Bouman, M. A., Bueno de Mesquita, A. E., Slappendel, S., 1978a. Perimetry of contrast detection thresholds of moving spatial sine wave patterns. I. The near peripheral visual field (eccentricity 0°-8°). Journal of the Optical Society of America 68 (6), 845–849.

Koenderink, J. J., Bouman, M. A., Bueno de Mesquita, A. E., Slappendel, S., 1978b. Perimetry of contrast detection thresholds of moving spatial sine wave patterns. II. The far peripheral visual field (eccentricity 0° - 50°). Journal of the Optical Society of America 68 (6), 850–854.

Koenderink, J. J., Bouman, M. A., Bueno de Mesquita, A. E., Slappendel, S., 1978c. Perimetry of contrast detection thresholds of moving spatial sine wave patterns. III. The target as a sensitivity controlling parameter. Journal of the Optical Society of America 68 (6), 854–60.

Koenderink, J. J., Bouman, M. A., Bueno de Mesquita, A. E., Slappendel, S., 1978d. Perimetry of contrast detection thresholds of moving spatial sine wave patterns. IV. The influence of mean retinal illuminance. Journal of the Optical Society of America 68 (6), 860–865.

Kontsevich, L. L., Tyler, C. W., 1999. Nonlinearities of near-threshold contrast transduction. Vision Research 39 (10), 1869–80.

Krantz, D. H., 1969. Threshold theories of signal detection. Psychological Review 76 (3), 308–324.

Kuffler, S. W., 1953. Discharge patterns and functional organization of mammalian retina. Journal of Neurophysiology 16, 37–68.

Kulikowski, J. J., Abadi, R., King-Smith, P. E., 1973. Orientational selectivity of grating and line detectors in human vision. Vision Research 13, 1479–1486.

Laming, D., 2013. Probability summation - a critique. Journal of the Optical Society of America A (in press).

Landy, M. S., Graham, N., 2004. Visual Perception of Texture. In: Chalupa, L. M., Werner, J. S. (Eds.), The Visual Neurosciences. MIT Press, Cambridge, MA, Ch. 9, pp. 1106–1118.

Lasley, D. J., Cohn, T. E., 1981. Why discrimination may be better than detection. Vision Research 21, 273–278.

Ledgeway, T., McGraw, P., Simmers, A., 2011. What is the spatial integration area for global motion perception in human central vision? Journal of Vision 11 (11), 704–704.

Legge, G. E., 1978. Space domain properties of a spatial frequency channel in human vision. Vision Research 18, 959–969.

Legge, G. E., Foley, J. M., 1980. Contrast masking in human vision. Journal of the Optical Society of America 70 (12), 1458–1471.

Levi, D. M., Harwerth, R. S., 1977. Spatio-temporal interactions in anisometropic and strabismic amblyopia. Investigative Opthalmology & Visual Science 16 (1), 90–95.

Levi, D. M., Harwerth, R. S., Manny, R. E., 1979. Suprathreshold spatial frequency detection and binocular interaction in strabismic and anisometropic amblyopia. Investigative Opthalmology & Visual Science 18 (7), 714–725.

Levi, D. M., Klein, S. A., 2008. What limits performance in the amblyopic visual system : Seeing signals in noise with an amblyopic brain. Journal of Vision 8:4 (1), 1–23.

Long, G. M., Tuck, J. P., 1991. Comparison of contrast sensitivity functions across three orientations: implications for theory and testing. Perception 20 (3), 373–380.

Lu, Z.-L., Dosher, B. A., 2008. Characterizing observers using external noise and observer models: assessing internal representations with external noise. Psychological Review 115 (1), 44–82.

Majaj, N. J., Carandini, M., Movshon, J. A., 2007. Motion integration by neurons in macaque MT is local, not global. The Journal of Neuroscience 27 (2), 366–370.

Mäkelä, P., Näsänen, R., Rovamo, J., Melmoth, D., 2001. Identification of facial images in peripheral vision. Vision Research 41, 599–610.

Manahilov, V., Simpson, W. A., 2001. Energy model for contrast detection: spatial-frequency and orientation selectivity in grating summation. Vision Research 41, 1547–60.

Manahilov, V., Simpson, W. A., McCulloch, D. L., 2001. Spatial summation of peripheral Gabor patches. Journal of the Optical Society of America A 18 (2), 273–282.

Marcum, J. I., 1947. A statistical theory of target detection by pulsed radar. Tech. rep., Rand Corporation, Santa Monica.

Marr, D., 1982. Vision. W. H. Freeman and Company, San Francisco.

Marr, D., Hildreth, E., 1980. Theory of Edge Detection. Proceedings of the Royal Society B 207 (1167), 187–217.

Marčelja, S., 1980. Mathematical description of the responses of simple cortical cells. Journal of the Optical Society of America 70 (11), 1297–1300.

Mayer, M., Tyler, C., 1986. Invariance of the slope of the psychometric function with spatial summation. Journal of the Optical Society of America A 3 (8), 1166–1172.

McKee, S. P., Levi, D. M., Movshon, J. A., 2003. The pattern of visual deficits in amblyopia. Journal of Vision 3, 380–405.

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), 1–21.

Meese, T. S., 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), 1–23.

Meese, T. S., Hess, R. F., 2007. Anisotropy for spatial summation of elongated patches of grating: a tale of two tails. Vision Research 47, 1880–1892.

Meese, T. S., Hess, R. F., Williams, C. B., 2005. Size matters, but not for everyone: Individual differences for contrast discrimination. Journal of Vision 5, 928–947.

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

Meese, T. S., Summers, R. J., 2009. Neuronal convergence in early contrast vision: binocular summation is followed by response nonlinearity and area summation. Journal of Vision 9:4 (7), 1–16.

Meese, T. S., 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), 1–28.

Melmoth, D. R., Rovamo, J. M., 2003. Scaling of letter size and contrast equalises perception across eccentricities and set sizes. Vision Research 43, 769–777.

Mishkin, M., Ungerleider, L. G., 1982. Contribution of striate inputs to the visuospatial functions of parieto-preoccipital cortex in monkeys. Behavioural Brain Research 6, 57–77.

Mullen, K. T., Beaudot, W. H. A., Ivanov, I. V., 2011. Evidence that global processing does not limit thresholds for RF shape discrimination. Journal of Vision 11:3 (6), 1–21.

Nachmias, J., 1981. On the psychometric function for contrast detection. Vision Research 21, 215–223.

Nachmias, J., Steinman, R. M., 1963. Study of Absolute Visual Detection by the Rating-Scale Method. Journal of the Optical Society of America 53 (10), 1206–1213.

Nelder, J. A., Mead, R., 1965. A simplex method for function minimization. The Computer Journal 7, 308–313.

Neri, P., 2002. Psychophysical reverse correlation as a potential bridge between perception and physiology. Perception 31 (ECVP Abstract Supplement).

Nichols, M. J., Newsome, W. T., 2002. Middle Temporal Visual Area Microstimulation Influences Veridical Judgements of Motion Direction. The Journal of Neuroscience 22 (21), 9530–9540.

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

Peirce, J. W., 2007. The potential importance of saturating and supersaturating contrast response functions in visual cortex. Journal of Vision 7:6 (13), 1–10.

Pelli, D. G., 1985. Uncertainty explains many aspects of visual contrast detection and discrimination. Journal of the Optical Society of America A 2 (9), 1508–1532.

Pelli, D. G., 1990. The quantum efficiency of vision. In: Vision: Coding and Efficiency. Cambridge University Press, Cambridge, pp. 3–24.

Pelli, D. G., 1991. Noise in the Visual System May Be Early. In: Landy, M., Movshon, J. A. (Eds.), Computational Models of Visual Processing. MIT Press, Cambridge, Ch. 11, pp. 147–152.

Perry, V. H., Cowey, A., 1985. The ganglion cell and cone distributions in the monkey's retina: Implications for central magnification factors. Vision Research 25 (12), 1795–1810.

Peterson, W. W., Birdsall, T. G., Fox, W. C., 1954. The theory of signal detectability. Information Theory, IRE Professional Group on 4 (4), 171–212.

Phillips, G. C., Wilson, H. R., 1984. Orientation bandwidths of spatial mechanisms measured by masking. Journal of the Optical Society of America A 1 (2), 226–232.

Pointer, J. S., Hess, R. F., 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., Norcia, a. M., 1998. Elongated physiological summation pools in the human visual cortex. Vision Research 38, 3735–3741.

Polat, U., Tyler, C. W. C. W., Mar. 1999. What pattern the eye sees best. Vision research 39 (5), 887–95.

Pöppel, E., Harvey, L. O., 1973. Light-Difference Threshold and Subjective Brightness in the Periphery of the Visual Field. Psychologische Forschung 36, 145–161.

Poynton, C., 1998. The rehabilitation of gamma. In: Rogowitz, B. E., Pappas, T. N. (Eds.), Human Vision and Electronic Imaging III, Proceedings of the SPIE/IS&T Conference 3299. Vol. 1998. SPIE, Bellingham, Washington, pp. 1–18.

Press, W. H., Flannery, B. P., Teukolsky, S. A., Vetterling, W. T., 1989. Downhill Simplex Method in Multidimensions. In: Numerical Recipes in Pascal. Cambridge University Press, Cambridge, UK, Ch. 10.4, pp. 326–330.

Prins, N., 2012. The psychometric function: the lapse rate revisited. Journal of Vision 12:6 (25), 1–16.

Prins, N., Kingdom, F. A. A., 2009. Palamedes: Matlab routines for analyzing psychophysical data. www.palamedestoolbox.org (accessed 07/01/13).

Quick, R. F., 1974. A vector-magnitude model of contrast detection. Kybernetik 16, 65–67.

Rijsdijk, J. P., Kroon, J. N., van der Wildt, G. J., 1980. Contrast sensitivity as a function of position on the retina. Vision Research 20 (3), 235–41.

Robson, J. G., Graham, N., 1981. Probability Summation and Regional Variation in Contrast Sensitivity Across the Visual Field. Vision Research 21, 409–18.

Rovamo, J., Franssila, R., Näsänen, R., 1992. Contrast sensitivity as a function of spatial frequency, viewing distance and eccentricity with and without spatial noise. Vision Research 32 (4), 631–637.

Rovamo, J., Luntinen, O., Näsänen, R., 1993. Modelling the Dependence of Contrast Sensitivity on Grating Area and Spatial Frequency. Vision Research 33 (18), 2773–2788.

Rovamo, J., Virsu, V., 1979. An estimation and application of the human corticial magnification factor. Experimental Brain Research 37, 495–510.

Rovamo, J., Virsu, V., Laurinen, P., Hyvärinen, L., 1982. Resolution of gratings oriented along and across meridians in peripheral vision. Investigative Opthalmology & Visual Science 23 (5), 666–670.

Rovamo, J., Virsu, V., Näsänen, R., 1978. Cortical magnification factor predicts the photopic contrast sensitivity of peripheral vision. Nature 271, 54–56.

Rovamo, J. M., Melmoth, D. R., 2002. Scaling of both grating size and contrast is necessary for equalising detection across eccentricities. Journal of Vision 2 (7), 200–200.

Sagi, D., 1990. Detection of an orientation singularity in Gabor textures: effect of signal density and spatial frequency. Vision Research 30 (9), 1377–1388.

Salzman, C. D., Newsome, W. T., 1994. Neural Mechanisms for Forming a Perceptual Decision. Science 264, 231–237.

Sasaki, Y., Rajimehr, R., Kim, B. W., Ekstrom, L. B., Vanduffel, W., Tootell, R. B. H., 2006. The radial bias: a different slant on visual orientation sensitivity in human and nonhuman primates. Neuron 51 (5), 661–670.

Shapley, R. M., Tolhurst, D. J., 1973. Edge detectors in human vision. Journal of Physiology 229, 165–183.

Simmers, A. J., Ledgeway, T., Hess, R. F., 2005. The influences of visibility and anomalous integration processes on the perception of global spatial form versus motion in human amblyopia. Vision Research 45, 449–460.

Simmers, A. J., Ledgeway, T., Hess, R. F., McGraw, P. V., 2003. Deficits to global motion processing in human amblyopia. Vision Research 43, 729–738.

Skottun, B. C., Bradley, A., Freeman, R. D., 1986. Orientation Discrimination in Amblyopia. Investigative Opthalmology & Visual Science 27 (4), 532–537.

Smith, E. L., Hung, L.-F., Harwerth, R. S., 2000. The Degree of Image Degradation and the Depth of Amblyopia. Investigative Opthalmology & Visual Science 41 (12), 3775–3781.

Snowden, R., Thompson, P., Troscianko, T., 2006. To the cortex. In: Basic Vision. Oxford University Press, Oxford, Ch. 3, pp. 67–95.

Strasburger, H., Rentschler, I., Jüttner, M., 2011. Peripheral vision and pattern recognition: a review. Journal of Vision 11:5 (13), 1–82.

Stromeyer, C. F., Klein, S., 1974. Spatial frequency channels in human vision as asymmetric (edge) mechanisms. Vision Research 14, 1409–1420.

Swets, J. A., 1961. Is there a sensory threshold? Science 134 (3473), 168–177.

Swets, J. A., Tanner, W. P., Birdsall, T. G., 1961. Decision processes in perception. Psychological Review 68 (5), 301–340.

Tanner, W. P., 1961. Physiological implications of psychophysical data. Annals of the New York Academy of Sciences 89, 752–765.

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

Thomas, J., 1978. Normal and amblyopic contrast sensitivity functions in central and peripheral retinas. Investigative Opthalmology & Visual Science 17 (8), 746–753.

To, M., Lovell, P. G., Troscianko, T., Tolhurst, D. J., 2008. Summation of perceptual cues in natural visual scenes. Proceedings of the Royal Society B 275, 2299–2308.

To, M. P. S., Baddeley, R. J., Troscianko, T., Tolhurst, D. J., 2011. A general rule for sensory cue summation: evidence from photographic, musical, phonetic and cross-modal stimuli. Proceedings of the Royal Society B 278, 1365–1372.

Tolhurst, D. J., 1972. On the possible existence of edge detector neurones in the human visual system. Vision Research 12, 797–804.

Tolhurst, D. J., Heeger, D. J., 1997. Comparison of contrast-normalization and threshold models of the responses of simple cells in cat striate cortex. Visual Neuroscience 14, 293–309.

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

Virsu, V., Rovamo, J., 1979. Visual resolution, contrast sensitivity, and the cortical magnification factor. Experimental Brain Research 37, 475–494.

Vorhees, H., Poggio, T., 1988. Computing texture boundaries from images. Nature 333 (26), 364–367.

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

Watamaniuk, S. N. J., 1993. Ideal observer for discrimination of the global direction of dynamic random-dot stimuli. Journal of the Optical Society of America A 10 (1), 16–28.

Watamaniuk, S. N. J., Sekuler, R., 1992. Temporal and spatial integration in dynamic random-dot stimuli. Vision Research 32 (12), 2341–2347.

Watson, A. B., 1987. Estimation of local spatial scale. Journal of the Optical Society of America A 4 (8), 1579–1582.

Watson, A. B., Ahumada, A. J., 2005. A standard model for foveal detection of spatial contrast. Journal of Vision 5, 717–740.

Webb, B. S., Ledgeway, T., McGraw, P. V., 2007. Cortical pooling algorithms for judging global motion direction. Proceedings of the National Academy of Sciences of the United States of America 104 (9), 3532–3537.

Webb, B. S., Ledgeway, T., McGraw, P. V., 2010. Relating spatial and temporal orientation pooling to population decoding solutions in human vision. Vision research 50, 2274–2283.

Wetherill, G. B., 1963. Sequential Estimation of Quantal Response Curves. Journal of the Royal Statistical Society. Series B (Methodological) 25 (1), 1–48.

Wetherill, G. B., Levitt, H., 1965. Sequential estimation of points on a psychometric function. The British Journal of Mathematical and Statistical Psychology 18 (1), 1–10.

Wichmann, F. A., Hill, N. J., 2001. The psychometric function: I. Fitting, sampling, and goodness of fit. Perception And Psychophysics 63 (8), 1293–1313.

Wilson, H. R., Wilkinson, F., Asaad, W., 1997. Concentric Orientation Summation in Human Form Vision. Vision Research 37 (17), 2325–2330.

Wilson, J. R., 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.

Wong, E. H., Levi, D. M., 2005. Second-order spatial summation in amblyopia. Vision Research 45 (21), 2799–2809.

Wright, M. J., Johnston, A., 1983. Spatiotemporal contrast sensitivity and visual field locus. Vision Research 23 (10), 983–989.

Appendices

APPENDIX A

Birdsall's theorem

A.1 Early noise and nonlinear transduction

A.1.1 Single-channel systems

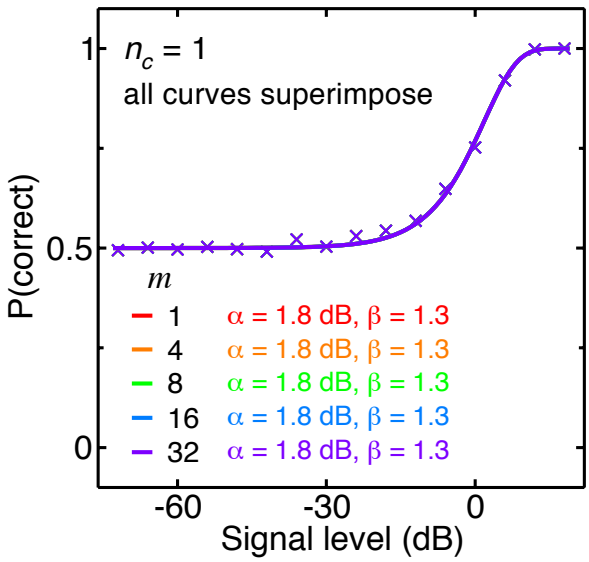

Figure A.1: Psychometric functions generated from a stochastic Monte Carlo simulation (5,000 iterations) of a detection task in a single-channel system ($n_c = 1$) featuring early Gaussian noise ($\mu = 0, \sigma = 1$) followed by nonlinear transduction with a range of different exponents (m).

The nonlinear transduction models tested in this thesis are all limited by additive internal noise placed after the transduction stage ("late" noise). Where the detection performance of a system with a single channel is limited by additive noise before monotonic nonlinear transduction ("early" noise) the noise has the effect of linearising the transducer. This means that for single-channel early noise models, the transducer exponent does not affect the trial-by-trial

performance (Lasley & Cohn, 1981). This can be illustrated by considering the effects of noise and nonlinear transduction on a single trial in a 2IFC task

$$R_{\text{target}} = [N(\mu, \sigma) + S]^m, \tag{A.1}$$

and

$$R_{\text{null}} = [N(\mu, \sigma)]^m, \tag{A.2}$$

where R_{target} is the response to the target interval, R_{null} is the response to the null interval, $N(\mu, \sigma)$ is a noise sample from a Gaussian distribution with a mean μ and standard deviation σ, S is the signal increment in the target interval, and m is the transducer exponent. The signal is detected on trials where

$$R_{\text{target}} > R_{\text{null}}. \tag{A.3}$$

Equations A.1 to A.3 show that when the dominant noise source is placed before the transducer then that transducer does not affect the trial-by-trial detection performance.

Figure A.1 shows the simulated behaviour of a single-channel system with early noise followed by nonlinear transduction. The simulations were conducted by modelling a system behaving according to Equations A.1 to A.3 with several different signal levels (S). Weibull psychometric functions were fitted to the simulated data using the Palamedes toolbox. A range of transducer exponents from $m = 1$ (a linear system) to $m = 32$ produce exactly the same behaviour. The threshold (Weibull α) remains constant at 1.8 dB (it is determined by the standard deviation of the simulated noise). The slope (Weibull β) of the psychometric functions is 1.3, which is characteristic of a linear system.

A.1.2 Multi-channel systems

In a system with multiple independently noisy channels, Birdsall's theorem no longer applies (Pelli, 1991). Figure A.2 shows the results of simulations of multi-channel systems with early noise followed by nonlinear transduction such that

$$R_{\text{target}} = \sum_{i=1}^{n_c} [N(\mu, \sigma) + S_i]^m, \tag{A.4}$$

and

$$R_{\text{null}} = \sum_{i=1}^{n_c} [N(\mu, \sigma)]^m, \tag{A.5}$$

where each $N(\mu, \sigma)$ is an independent noise sample, S_i is the signal level in each channel, and n_c is the number of channels. In this case, the different transducers do affect the behaviour of the systems. Increasing the transducer exponent in a range from $m = 1$ (a linear system) to $m = 32$ raises the threshold (Weibull α), and slightly shallows the slope (Weibull β). Increasing

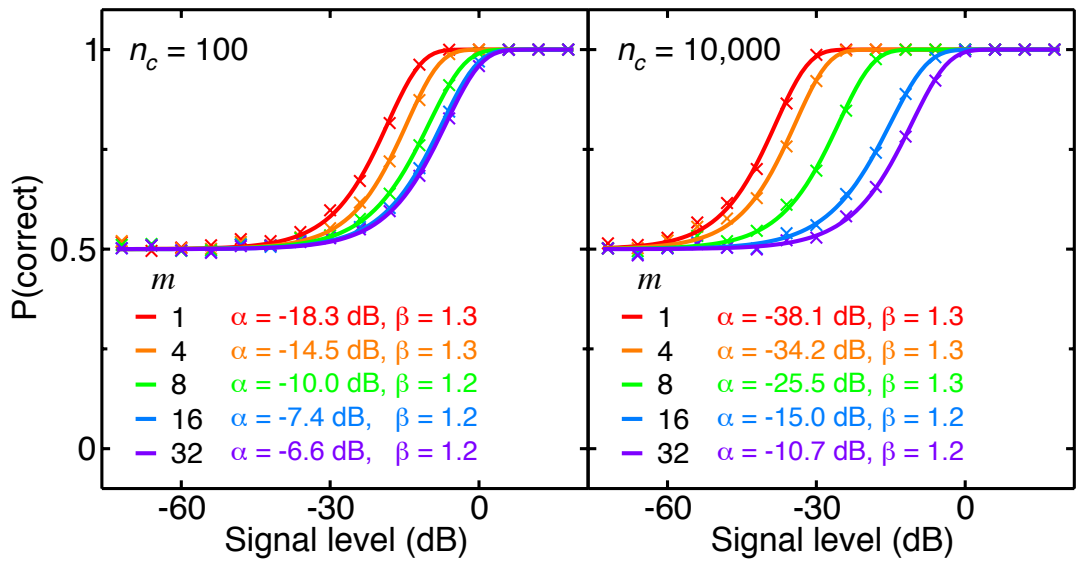

Figure A.2: Psychometric functions generated from stochastic Monte Carlo simulations (5,000 iterations) of a detection task in systems with 100 and 10,000 channels featuring independent early Gaussian noise ($\mu = 0, \sigma = 1$) in each channel followed by nonlinear transducers with different exponents (m).

the number of channels in the system (n_c) exaggerates the effect of the transducer exponent on the threshold (and also increases the overall sensitivity).

A.1.3 Area summation with early noise

Area summation behaviour in multi-channel systems would also differ from that expected if Birdsall's theorem was to apply. Figure A.3 shows the decreases in threshold that result from a doubling in the number of signal-carrying channels (from $\frac{n_c}{2}$ to n_c) in systems with different numbers of channels and with a range of transducer exponents. The simulated system monitored its channels ideally, attending only to relevant channels on each trial (ignoring the input from the non-signal channels in the half-signal condition). If Birdsall's theorem was to apply, the summation ratio would be expected to remain constant at $\sqrt{2}$ (see Section 5.3.4). Instead, it decreases as the exponent (m) increases. In the limit case this behaviour is expected, because Equations A.4 and A.5 will behave more like max operators as m tends to infinity. Simulating higher exponents would therefore lead to the early noise prediction asymptoting at the summation ratio predicted by a late noise exponent of 8 (due to the combination of the square law effect from the ideal summation strategy and a Minkowski exponent of 4 from the SDT probability summation model; see Section 5.4.3). The predictions in Figure A.3 show that as the number of channels (n_c) increases larger exponents are required to switch the behaviour away from that expected in a linear system.

As the psychometric function slopes in this model are still very shallow compared to the em-

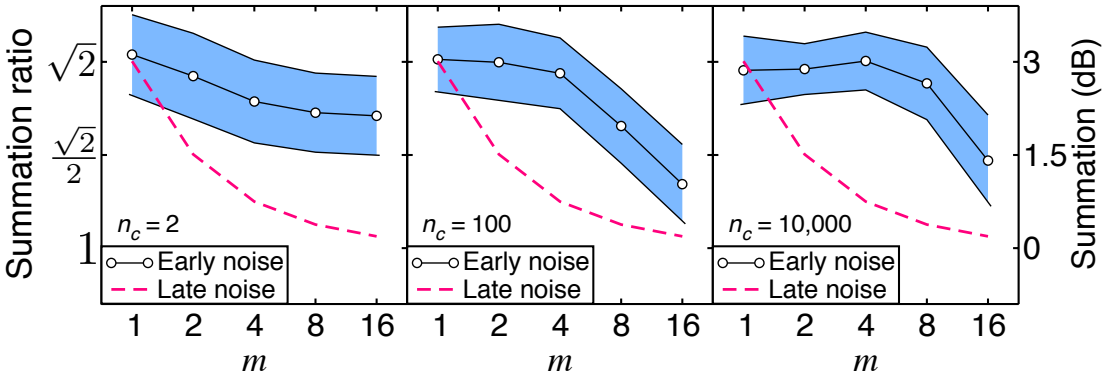

Figure A.3: Summation behaviour in multi-channel systems with either "early" noise placed *before* nonlinear transduction, or "late" noise placed *after* nonlinear transduction. The early noise predictions are calculated from the mean of 20 stochastic Monte Carlo simulations (500 iterations each). The blue shaded area shows the standard deviation of the simulated data. Late noise predictions are derived using Equation 5.22. Models featured either 2, 100, or 10,000 channels, and a range of transducer exponents from $m = 1$ to $m = 16$. The summation ratio plotted is that between thresholds for stimuli presented to half of the channels or all of the channels in the system. The model featured a matched template, so the system monitored only relevant inputs (ignoring non-signal channels).

pirical data, a complete summation model based on early noise would have to rely on an additional component to steepen them. One way to achieve this would be to introduce a large amount of uncertainty about which channels contain task-relevant signal (Pelli, 1985; Tyler & Chen, 2000). Such large amounts of intrinsic uncertainty also predict there should be little effect on detection threshold for interleaving stimuli of different sizes versus blocking those stimuli. Results from studies that have made this manipulation however do show differences in threshold across these two conditions (Meese & Summers, 2012). In the absence of another method to reconcile this disparity between the predictions of the early noise model and empirical findings, I conclude that it is not necessary to consider such models in this thesis.

APPENDIX B

MATLAB code

B.1 Log-Gabors

B.1.1 MATLAB code to produce log-Gabor patches

The function supplied below generates cartesian-separable log-Gabor patches, as described in Section 3.4.5. It has been tested in MATLAB versions 2010a and 2012a.

B.1.2 loggabor.m

```
function logGabor = loggabor(w, pixPerDeg, sf, ori, sfBW, oriBW, phase, c)
% ASB - January 2012 - adapted and optimised from code by Tim Meese
% Makes a log-Gabor stimulus (cartesian-separable)
% INPUT:  w         = size of the output image (pixels)
%         pixPerDeg = number of pixels in one degree of visual angle
%         sf        = spatial frequency (c/deg)
%         ori       = orientation (degrees)
%         sfBW      = sf bandwidth (octaves, full width @ half height)
%         oriBW     = ori bandwidth (degrees, half width @ half height)
%         phase     = phase (degrees)
%         c         = nominal contrast (0-1)
% OUTPUT: lgImage   = an image scaled between -1 and +1

    % Convert degrees to radians.
    ori = ori*(pi/180);
    oriBW = oriBW*(pi/180);
    phase = phase*(pi/180);
    % SF scaled to Fourier dom & w.
    fFT = sf * (w / pixPerDeg);
```

```
    % CALC OF LOG-GAUSSIAN SPREAD PARAMETERS
    % N.B: 1.18sigmaU = 0.5W = 0.5SFBandwidth.
    logSigmaU = log10(2^sfBW) / (2 * 1.18); % units c/deg
    lgV       = log10(cos(oriBW));
    gSF       = exp(-(lgV^2 / (2*logSigmaU^2)));

    % Multiplying by SF maintains scale invariance.
    sigmaV = 2 * sf * sin(oriBW) * sqrt((-1)/(2*log(0.5/gSF)));

    % Absolute value of sigmaV
    % (which is then converted to F-domain pixels).
    sigmaVPix = abs(sigmaV) * w/(pixPerDeg * 2);

    % Preallocate for speed
    cxGauss1 = ones(w) .* complex(0,0);
    cxGauss2 = cxGauss1;
    % Fourier origin = [1,1] px
    u0 = ((w+2)/2); v0 = ((w+2)/2);

    % MATRIX OPERATIONS (replaces loop)
    u = meshgrid(1:w,1:w); v = u';
    % Polar coordinates
    r = sqrt((u-u0).^2 + (v-v0).^2);
    % Rotate coordinates to desired orientation
    ang = atan2((v-v0),(u-u0)) - ori;
    % convert to Cartesian coordinates
    uft = r.*cos(ang); vft = r.*sin(ang);

    % +uft is +ve :: Equiv of uft - fFT (Linear Gabor)
    lgf = log10(+uft(uft>0) ./ fFT);
    gauss = exp(-(lgf.^2)./(2*logSigmaU^2)) .* ...
            exp(-(vft(uft>0).^2) * 1/(2*sigmaVPix^2));
    cxGauss1(uft>0) = complex(gauss.*sin(phase), -gauss.*cos(phase));

    % -uft is +ve :: Equiv of uft - fFT (Linear Gabor)
    lgf = log10(-uft(uft<0) ./ fFT);
    gauss = exp(-(lgf.^2)./(2*logSigmaU^2)) .* ...
            exp(-(vft(uft<0).^2) * 1/(2*sigmaVPix^2));
    cxGauss2(uft<0) = complex(gauss.*sin(phase), gauss.*cos(phase));
    % Combine the gaussians to produce the filter
    cxGauss = cxGauss1 + cxGauss2;

    % transform back to spatial domain
    imageRaw = real(ifftshift(ifft2(fftshift(cxGauss))));
    % scale by peak (don't trust Michelson)
    logGabor = c .* (imageRaw ./ max(abs(imageRaw(:))));

return
```

B.2 The witch's hat

B.2.1 MATLAB code to produce a witch's hat attenuation surface

The function supplied below generates a witch's hat attenuation surface (see Section 4.7.4). This code has been tested in MATLAB versions 2010a and 2012a.

B.2.2 witchhat.m

```
function witchHat = witchhat(imSize,pixPerCycle,p)
% ASB - February 2012 - adapted and optimised from code by Tim Meese
% Makes a witch's hat attenuation surface
% INPUT:  imSize - the output image size
%         pixPerCycle - the wavelength of the spatial frequency
%         p - a vector describing the bilinear falloff
%         [m1_sup m2_sup m1_inf m2_inf m1_hor m2_hor kneepoint]
% OUTPUT: witchHat - the witch hat attenuation surface

        % Convert input p to full 12-parameter witch's hat description
        x1dBL  = p(5);  x1dBR  = p(5); y1dBU  = p(1); y1dBD  = p(3);
        x2dBL  = p(6);  x2dBR  = p(6); y2dBU  = p(2); y2dBD  = p(4);
        kinkxL = p(7);  kinkxR = p(7); kinkyU = p(7); kinkyD = p(7);

        % Create variables holding position information
        x0 = ((imSize+1)/2); y0 = ((imSize+1)/2);
        [xx, yy] = meshgrid(1:imSize, 1:imSize);
        x1 = xx-x0; y1 = yy-y0;

        % Do geometry
        x1sq = real(x1.*x1); y1sq = real(y1.*y1);
        r = real(sqrt(x1sq + y1sq));
        ang = atan2(y1, x1);

        % Populate H&V meridian column vectors with parameters
        m1h   = x1dBL.*(x1<0)  + x1dBR.*(x1>0);
        m2h   = x2dBL.*(x1<0)  + x2dBR.*(x1>0);
        m1v   = y1dBU.*(y1>0)  + y1dBD.*(y1<0);
        m2v   = y2dBU.*(y1>0)  + y2dBD.*(y1<0);
        kneeh = kinkxL.*(x1<0) + kinkxR.*(x1>0);
        kneev = kinkyU.*(y1>0) + kinkyD.*(y1<0);

        % Interpolate between cardinals using the equation for an ellipse
        % m1 & m2 (N.B. using the reciprocal of the gradient parameters)
        m1dB = (1./(m1h.*m1v))./(sqrt((cos(ang)./m1v).^2 + ...
                                      (sin(ang)./m1h).^2));
        m2dB = (1./(m2h.*m2v))./(sqrt((cos(ang)./m2v).^2 + ...
```

```
                                        (sin(ang)./m2h).^2));
        % kneepoint (N.B. not using the reciprocal of the knee parameter)
        knee = (kneeh.*kneev)./(sqrt((kneev.*cos(ang)).^2 + ...
                                        (kneeh.*sin(ang)).^2));

        % This is the standard bilinear equation, as a function of radial
        % distance. Three parameters (m1, m2 and the knee) vary with angle.
        % Note that the gradient parameters (m1 and m2) are stored as
        % reciprocals, so must be reciprocated again in this equation.

        k1 = log10(1 ./ (10 .^ ((1./m1dB - 1./m2dB) .* knee) + 1));

        surface = -log10( ...
                        (10.^((1./m1dB) .* r./pixPerCycle)) ./          ...
                        (10.^((1./m1dB - 1./m2dB) .* knee) +            ...
                        (10.^((1./m1dB - 1./m2dB) .* r./pixPerCycle))) ...

        witchHat = flipud(surface);

return
```

APPENDIX C

Amblyope subjects

C.1 Table of amblyope subject information

		Corrected acuity	
#	Type	Right eye	Left eye
1	Strabismic	20/20	20/400
2	Strabismic/Anisometropic	20/250	20/20
3	Strabismic	20/200	20/30
4	Strabismic/Anisometropic	20/20	20/60
5	Strabismic	20/20	20/100
6	Strabismic	20/100	20/20

Table C.1: Details for the six amblyope observers tested in Chapter 8.