2013

Estimation of the parameters of a boundary contour system using psychophysical hyperacuity experiments

BOSTON UNIVERSITY
GRADUATE SCHOOL OF ARTS AND SCIENCES

Dissertation

ESTIMATION OF THE PARAMETERS OF A BOUNDARY CONTOUR SYSTEM USING PSYCHOPHYSICAL HYPERACUITY EXPERIMENTS

by

HARALD RUDA

A.B., Harvard University, 1988
M.A., Boston University, 2011

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
2013

Approved by

First Reader [redacted]

Ennio Mingolla, Ph.D.
Professor of Psychology

Second Reader [redacted]

Arash Yazdanbakhsh, Ph.D. MD
Research Assistant Professor of Cognitive & Neural Systems

Third Reader [redacted]

Takeo Watanabe, Ph.D.
Professor of Psychology

Acknowledgements

Who should be thanked? In short, everyone! And especially if you are reading this, I thank you! You are one among a very small elect group indeed. There are many who have been generous with their time and ideas.

John Daugman was a strong presence in my undergraduate years; he was the first to teach me about vision and I found it compelling. He also set in my mind forever the idea that there can be no understanding without theory. He was also responsible for telling me about this new graduate program at Boston University.

Steve Strogatz was a resident at my undergraduate house (Winthrop) and was also one of the teachers for an early math course. I fondly remember his weekly Science Table.

Steve Grossberg of course founded the Cognitive and Neural Systems Program. That would be enough to be included in this list, but his ideas have resonated (!) with me. His writings and teachings have had, and still have, a monumental effect.

Ennio Mingolla is the one teacher who has inspire me the most. His method is always to ask questions and to probe deeply, to always go below the surface. As my major adviser, I am especially grateful that he has been willing to see this project to conclusion.

James Todd taught me about visual psychophysics, and showed me the mind of an experimentalist and how the right experiments can explain so much.

Takeo Watanabe is another teacher who has been an inspiration. I am grateful that he has agreed to be a reader for this dissertation.

Associate Dean W. Jeffrey Hughes is probably not thanked very often, at least no in his role as dean. However he provided the final incentive to actually finish this project, and for this he will never be forgotten.

Arash Yazdanbakhsh has been singularly influential in the finishing phase of this project. In addition to his valuable intellectual contributions, he has been a constant “cheerleader”, and he also agreed to be a reader.

The other professors and staff at the department has always been extremely professional and helpful, I thank them for all their help and hard work.

The other students and post-docs, always an interesting group have been helpful. In addition the lunch table of the last couple of years has been a source of many stimulating discussions and good ideas.

I also thank my family, my loving wife Tammie, and my daughters Katarina and Annika, they never questioned my commitment to the project. My parents and my brothers have also been supportive, and it looks like I have won a bet made with my brothers in September 2010.

And of course, last but not least, there is also the financial support. It is only fair to acknowledge that this work was supported in part by British Petroleum (89A-1204), DARPA (AFOSR 90-0083), and AFOSR (90-0175). However, the majority of financial, and also emotional, support was provided by Tammie L. Ruda.

ESTIMATION OF THE PARAMETERS OF A BOUNDARY CONTOUR SYSTEM USING PSYCHOPHYSICAL HYPERACUITY EXPERIMENTS

(Order No. 3529062 from UMI)

HARALD RUDA

Boston University Graduate School of Arts and Sciences, 2013

Major Professor: Ennio Mingolla, Professor of Psychology

ABSTRACT

Visual hyperacuity enables observers to make accurate judgments of the relative positions of stimuli when the differences are smaller than the size of a single cone in the fovea. Because hyperacuity can serve as a gauge for precisely measuring characteristics of the visual system, it can provide stringent tests for models of the visual system. A variant of the Boundary Contour System (BCS) model is here used to clarify previously unexplained psychophysical hyperacuity results involving contrast polarity, stimulus separation, and sinusoidal masking gratings. Two-dot alignment thresholds were studied by Levi & Waugh (1996) by varying the gap between the dots, with same and opposite contrast polarity with respect to the background, and also with and without band-limited sinusoidal grating masks of different orientations. They found that when the gap between the dots is small (6 arcmin), different patterns of misalignment thresholds are obtained for the same and different contrast polarity conditions. How-

ever, when the gap is large (24 arcmin), the same pattern of thresholds was obtained irrespective of contrast polarity.

The simulations presented here replicate these findings, producing the same pattern of results when varying the gap between the dots, with same and opposite contrast polarity with respect to the background, and also with and without sinusoidal grating masks of different orientations. The vision model used (BCS) is able to produce these patterns because of its inherent processing using contrast insensitivity, spatial and oriented competition, and long-range completion layers. A novel aspect of the model is the use of sampled field processing, which simplifies the model's equations. Modified Hebbian learning and a neural decision module are proposed as mechanisms that link the vision model's outputs to a decision criterion. All model parts have plausible neurobiological correlates.

In addition, psychophysical hyperacuity experiments served to map the limits of inhibitory spatial interactions. The results show that inhibition occurs even when only half of the split flanking line of Badcock & Westheimer (1985b) is used, suggesting that subthreshold activity in units representing the line extends beyond the end of the line. Furthermore, strong inhibition was observed with a flanking illusory line grating.

Preface

If I communicate anything in this dissertation, I hope it is something about how the human visual system works. I find it a fascinating topic. The particular sub-topic chosen here, hyperacuity, provides an especially clear window into the understanding of how it is why we see what we see. I hope this window is clear and visible to at least a few others as well.

This dissertation was not written overnight. It has taken longer than the Big Dig. It may be over budget, but not by billions of dollars. But the question I often ask myself is, why did it take so long? There are many parts to this answer, but an interesting take might be that the linking hypothesis for connecting the model output to observed responses was not ready yet when the project was started. I knew that something was needed, but I did not know what, exactly. When I was introduced to the work of the Dosher lab (Liu *et al.*, 2010; Petrov *et al.*, 2005) pieces started to fall into place, and the work could finally be finished.

Table of Contents

List of Tables

List of Tables

List of Abbreviations

′	arcmin, minute of arc, one 60th of a degree of visual angle
″	arcsec, seconds of arc, one 60th of an arcmin
2AFC	Two Alternative Forced Choice
3AFC	Three Alternative Forced Choice
4AFC	Four Alternative Forced Choice
ALE	Adaptive Logistic Estimation
APE	Adaptive Probit Estimation
ART	Adaptive Resonance Theory
ART2	Adaptive Resonance Theory, for analog inputs
ART2A	Adaptive Resonance Theory, for analog inputs, algorithmic
BCS	Boundary Contour System
CC Loop	Competitive-Cooperative Loop
cpd	Cycles per degree, the frequency component of spatial frequency
CRT	Cathode Ray Tube
FCS	Feature Contour System
jnd	Just Noticeable Difference
L1	Unusual mathematical measure of length, "city-block" metric
L2	Standard mathematical measure of length, Euclidean distance
LGN	Lateral Geniculate Nucleus
MCS	Method of Constant Stimuli
MLE	Maximum Likelihood Estimation
ODE	Ordinary Differential Equation
PEST	Parameter Estimation by Sequential Testing
PSE	Point of Subjective Equality
QUEST	A Bayesian adaptive psychometric method, related to PEST
RT	Response Time
SDT	Signal Detection Theory
STM	Short Term Memory

Chapter 1:
Introduction

1.1 The Long History of Hyperacuity

Pierre Vernier (1580-1637) realized that the ability of humans to detect very slight offsets of abutting lines could be used to make very precise measurements. He devised the scale which is nowadays commonly found on equipment requiring precise visual judgement. In the 19th century when scientists discovered the composition of the retina, they also discovered the paradox that the spacing of receptors in the fovea is many times larger than the minimum detectable Vernier offset (Volkmann, 1863; Wülfing, 1892). Westheimer (1975) coined the term *hyperacuity* to describe the ability of subjects to discriminate pattern differences smaller than the diameter of a foveal cone. Even today, there is no general agreement about the mechanism used to accomplish this feat.

Research in the area since the time of Hering (1899) and Best (1900) has been scattered (French, 1920; Berry, 1948; Ludvigh, 1953). Hering (1899) postulated that the visual system performs summation along the lines in the display to achieve a higher signal-to-noise ratio. This explanation was refined (Weymouth *et al.*, 1923; Andrews *et al.*, 1973) and is still the explanation taught in undergraduate courses. Ludvigh (1953) showed that Hering's explanation cannot be true, because he measured the same low threshold using two dots; and since dots have no edges, there can be no integration of information along the edge. Westheimer revived interest in hyperacuity with a series of experiments reported in the 1970's (Westheimer & Hauske, 1975; Westheimer & McKee, 1977a, 1977b); and a large number of experiments have been reported since then. New theories which try to explain hyperacuity have also been postulated, mainly theories based on frequency selective spatial filters. Some theories also include line-element processing from explanations of color vision (Wilson & Gelb, 1984).

An explanation for hyperacuity ought to be related to a more general theory of early visual processing. Although much has been learned about the physiology of the visual cortex in recent years, there is still no widely accepted theory of vision which explains why the different cell types in visual cortex exist, and explains their precise function. Marr (1982) proposed a general theory which is well known, but it presents no direct physiological predictions because it abstracts the algorithm from the computational medium. However, many current theories which attempt to explain some part of visual function tend to include the simple and complex cells of cortical area V1, with fairly accurate descriptions of their receptive fields. Thus a modern explanation of *hyperacuity* should also include facets of well known physiology such as the known characteristics of cells in V1 and V2.

1.2 Optical Considerations

One may think that it is not so surprising that humans can detect offsets much smaller than the diameter of a foveal cone; after all, the physical stimulus is projected onto the retina and if there is a difference between two stimuli, then some clever processing should reveal it. The situation is not quite as simple as that however, because the physical stimulus is degraded by the optical apparatus before it reaches the retina. Any practical lens[1] blurs the image passing through it. This blurring is called a point-spread function. For a normal eye with 2 mm pupil the point spread function has the shape similar to a narrow Gaussian, and its width at half-height is about 1′ (Westheimer, 1979, 1981). A cone in the fovea has a diameter of about 30″, and is therefore fairly closely matched to the point-spread function.[2]

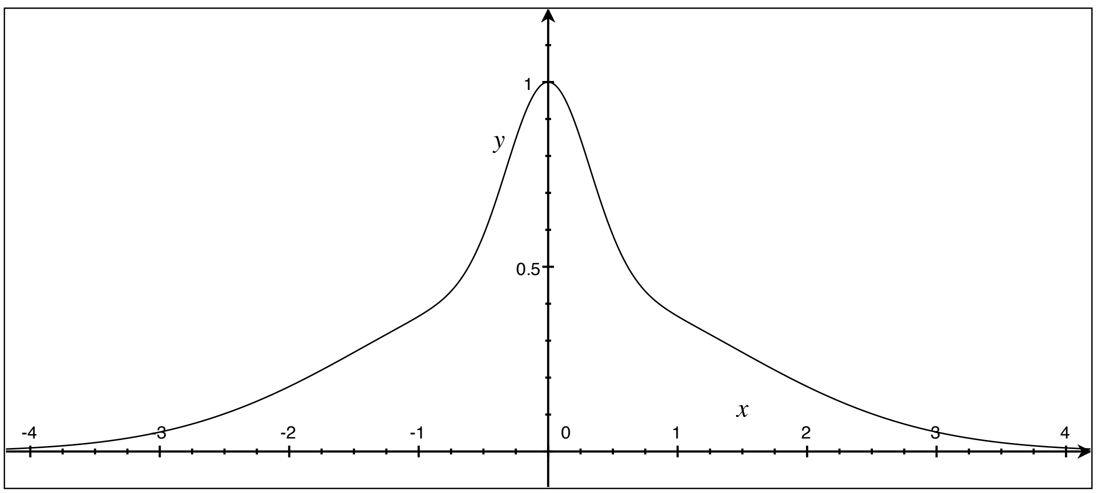


Figure 1-1: Line-spread function of the eye for a 2 mm pupil. The figure is replotted from Geisler (1984) which is based on data from Campbell & Gubisch (1966). Note that the width of the curve at half-height is slightly more than one arcmin, more than the combined diameters of two foveal cones!

The point-spread function is the major determinant of ideal visual acuity. Regular visual acuity is measured with a two-dot discrimination task; i.e., what is the minimum separation of two dots which still yields a percept of two dots. This type of acuity can be explained using point-spread functions and minimum

1. That is, a lens with a larger than infinitesimal diameter. Only a pinhole lens perfectly reproduces the image. A practical lens functions as a low-pass filter and therefore distorts the signal. A larger lens is desirable however, because it allows more photons through. As any photographer knows, there is an inherent trade-off between sharpness and amount of light.
2. The symbols ′ and ″ will be used to denote minutes and seconds of arc respectively.

detectable contrast (Westheimer, 1979). This situation is shown in Figure 1-2. The combination of the two point spread functions must have a "valley" with a contrast which is detectable. When this is calculated one gets approximately the same threshold as is measured for two dot discrimination, around 1′. Since hyperacuity involves relative shifts in the location of point-spread functions, it is not subject to this 1′ limit. What is required is the ability to localize stimuli which have been degraded by the point-spread function (for a thin line it may be called the line-spread function). The processing that produces hyperacuity must be quite sophisticated.

Figure 1-2: The contrast "valley" formed by the point spread functions of two nearby dots. Deciding whether the stimulus consists of one or two dots is simply a matter of being able to see the dip in contrast between the two dots.

1.3 The Hyperacuity Phenomenon

Hyperacuity has been studied most extensively in the Vernier acuity paradigm; some examples of which are shown in Figure 1-3. Two stimulus elements which are almost aligned in some direction (say vertical, as in Figure 1-3) are presented simultaneously; the task of the subject is to specify in which direction one of the elements (say the top element) has been shifted. The elements used are often lines or dots and a variety of parameters may be modified, such as the line length, separation, and contrast. Remarkably, the thresholds[3] for detecting the direction of shift are 5-6″ under optimal conditions. This is equivalent to detecting a shift of only 2.5 mm at a distance of 100 m! In non-metric units this becomes a tenth of an inch at a distance of 110 yards. This is remarkable because, as mentioned before, this is a fraction of the size and spacing of retinal photoreceptors in the fovea. Vernier acuity is only one of many tasks in which subjects display hyperacute performance.

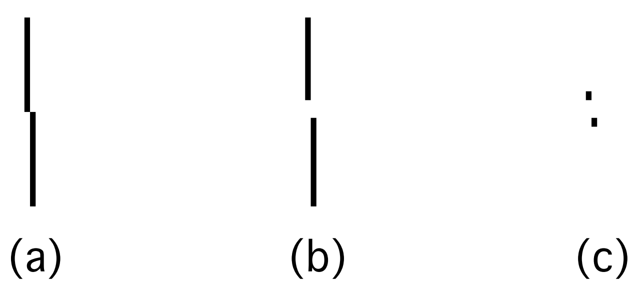


Figure 1-3: Some Vernier acuity stimuli: (a) Shows the traditional Vernier stimulus with thin abutting lines. (b) Introduces a slight separation between the line ends. (c) Depicts the result of shortening the lines to dots.

It is interesting to compare Vernier acuity with displacement acuity. In the displacement acuity paradigm the reference and test objects are not presented simultaneously but sequentially. For example: one may present to the subject a thin vertical line for 1 second and halfway during the presentation the line is shifted slightly to the left or to the right. By using selected displacements it is possible to compute the psychometric function and find the threshold for detecting the direction of displacement (here left or right). In normal subjects this threshold is about 10-12″. Furthermore, the size and shape of the stimulus or the

3. In this dissertation, “thresholds” will always refer to the 75% correct thresholds.

direction in which it is shifted is irrelevant and leaves the threshold unchanged (Westheimer, 1979). Even modifying the spatial frequency content by using sinusoidal gratings instead of lines has no effect on the displacement threshold. Westheimer (1978) found a constant retinal displacement threshold of 10-12″ for spatial frequencies from 3 to 25 cycles/degree.

Table 1-1: Anatomical, Physiological, and Psychophysical Data Relevant to Hyperacuity Phenomena (Human data except where noted; thresholds refer to 75% correct performance)

Point-spread function of eye (width at half-height)	~1' (Westheimer, 1979)
Foveal cone diameter (also spacing, packed)	~30" (Westheimer, 1981)
Size of smallest cortical receptive fields (macaque)	2' (Poggio *et al.*, 1977)
Orientation bandwidth of cortical cells (macaque)	~15° (Koenderink, 1984)
Acuity (2-point discrimination)	~1' (Westheimer, 1981)
Displacement threshold (for all stimuli)	10-12" (Westheimer, 1978)
Vernier threshold	5-6" (Westheimer, 1979)
Bisection threshold	3-4" (Klein & Levi, 1985)
Bisection threshold (optimal) (brightness, not displacement judgment)	0.85" (Klein & Levi, 1985)
Orientation discrimination (for lines 30' long)	0.17° tilt (Westheimer, 1981)
Stereo hyperacuity (displacement yielding jnd in depth)	3" (Westheimer, 1979)

There are other types of phenomena which exhibit hyperacuity. Bisection acuity, for example, can be as good as 1″; this is not a pure hyperacuity judgement however. Table 1-1 provides a summary of some data relevant to hyperacuity phenomena.

1.4 Local Sign and Filter Theories

Hering (1899) proposed the idea that information is averaged over the length of the lines in the Vernier display. This idea has developed into what is now called the local-sign theory. The original idea was that every photo-receptor on the retina codes a location in the visual field, namely its location. When light hits this photo-receptor, it sends a signal to the brain coding the light-level at that location in visual space. When a thin line is presented in visual space many photo-receptors become active, some of them partially.

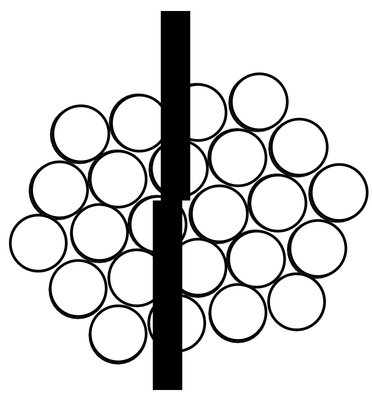


Figure 1-4: Diagram of a Vernier stimulus projected onto photo-receptors in the fovea. Looking at the responses of all the photo-receptors in the vicinity of the lines yields a much more accurate estimate of the position of the lines than just choosing one photo-receptor for each line. Note that the blurring effects of the line-spread function have been omitted in this figure.

By averaging the signals from photo-receptors along the length of the lines a very accurate estimate of the position of each line can be deduced. Although modern versions of the local-sign theory do not demand that the photo-receptor signals are passed unmodified to the cortex (as they clearly are not), the emphasis is still on positional coding. Another way to describe the hypothesis is that first the positions of the lines are established, and then those positions are compared. Clearly, the positions have to be known with very high accuracy for this to be feasible in the case of hyperacuity.

Westheimer (1981) proposed a variation of this theory which is best described as the *modular* local-sign theory. He suggested that there are spatial modules operating in the visual field, and that each module has the ability to accurately calculate the position of one stimulus element. If more than one element is present within the module however, the positions will not be nearly as accurate. This modification of the local-sign theory was necessary to account for some of the spatial interference data described by Westheimer & Hauske (1975).

Filter theories on the other hand try to explain the hyperacute judgement of *relative* positions without knowing the precise location of individual elements. This is done by having a number of spatial filters of varying spatial frequency and orientation preference respond to the stimulus. The responses of the filters is then usually compared with some accepted standard to determine if there is a hyperacute displacement. In other words, a difference is detected if the pattern of responses deviates from the ideal stimulus; and this difference can be hyperacute. The precise location of stimulus elements is not explicitly coded in filter theories.

Wilson (1986) successfully applies a filter model to a number of reported hyperacuity results. However, there remain a number of stimulus paradigms where additional mechanisms are required to explain the pattern of the data. Some of the mechanisms which are necessary are competition among spatial locations, long-range boundary completions, and also feedback between stages. A filter bank with these stages added onto it becomes very similar to the

Boundary Contour System (BCS) first described by Grossberg & Mingolla (1985a, 1985b).

This dissertation seeks to find a unified mechanism to explain all the tasks where hyperacuity thresholds have been reported. Since there are both theoretical and experimental reasons to believe that something similar to the BCS is part of the visual system (Grossberg & Mingolla, 1985a, 1985b), the BCS has been chosen as an appropriate model to use. It has the ability to explain almost all hyperacuity phenomena. The problem with using a complex model such as the BCS is that there are a large number of free parameters, and values of these parameters appropriate for the workings of the human visual system have not yet been reported. The intent here is to use a selected part of the hyperacuity database in an attempt to find suitable parameters; and then use these parameters to model the rest of the hyperacuity data.

The full range of experimental data necessary to constrain the BCS model is not yet available in the literature. Some of the experiments which would be most useful for determining the parameters of the competitive and cooperative mechanisms have simply not been performed. Occasionally the experimental data is not useable because there has been confusion about how to analyze the data. For these reasons, a number of psychophysical experiments[4] were performed in order to fill in the missing data. To determine the final parameters of the model a number of computer simulations were also performed. Therefore, this dissertation reports data from experiments as well as simulations.

---

4. Henceforth simply *experiments*, in contrast with computer *simulations*.

1.5 Hyperacuity as a Measuring Device

The very high sensitivity of hyperacuity phenomena makes it possible to use experimental data from hyperacuity tasks to fit models of early visual perception to within small tolerances. If the BCS model is close to being a viable model, then it should be possible to fit the model to a subset of the experimental data available. If the fitting works, then the derived parameters will be true not only for the experimental data used, but for all experimental data (within observer and experimental variability). The model with the derived parameters can then be evaluated by applying it to other experimental situations.

This dissertation uses physiological and psychophysical data in an attempt to put bounds on the otherwise elusive parameters of the BCS as they would be in the human visual system. Suitable parameters are estimated by focusing on Vernier and displacement acuity tasks. Recognizing that there are some key data for the determination of these parameters which have not yet been collected, a small number of parametric psychophysical experiments are also reported to this end. To test whether the parameters selected are related to those in the human visual system, the model is applied to other hyperacuity situations which have not been previously explained.

1.6 Organization of this Dissertation

This dissertation proceeds with a review of some hyperacuity literature and modeling efforts in Chapter 2. This is followed by a report on original hyperacuity experiments in Chapter 3. This chapter is a modified version of a previously published paper by Ruda (1998), that maps out the full extent of the inhibitory zones, as well as test some other ideas. Some work at simplifying the BCS model description and initial parameter settings is explained in Chapter 4. Chapter 5 explains the model and presents the modeling results. Chapter 6 is left for conclusions and ruminations about possible future work. Chapters 3 and 5 contain the main results of this work, and are structured in such a way that they can be read independently of the other chapters.

Chapter 2:
Hyperacuity Data & Models

This chapter discusses a large number of studies involving hyperacuity. It also discusses existing models of hyperacuity.

2.1 Vernier Acuity Literature

The previous chapter summarily dismissed Hering's idea of averaging local-signs to produce a hyperacute judgement. While this dismissal was justified, related ideas about temporal averaging are more difficult to ignore. Westheimer & McKee (1977a) showed that the visual system is capable of surprising feats of temporal averaging. For example, they showed that the visual system has a temporal integration period of about 20 ms for Vernier acuity displays. Longer delays between short presentations of the lines lead to increased thresholds. More interestingly, Westheimer & McKee found that moving the Vernier display across the retina does not interfere with hyperacuity judgements. That is, even though the Vernier lines are never stationary, correlations are built up over time which allows the task to be performed at hyperacuity. Westheimer & McKee (1977a) showed that the region over which such integration takes place is about 4′ wide. This is really a very odd result which does not seem easily explainable by a filter-type theory. It seems that it will be necessary to include elements of time-dependent processing, perhaps even explicit modeling of motion, into any theory aimed at explaining this data.

Westheimer (1981) implicitly made a prediction with his modular modification of the local-sign theory that the judgement of relative location within a module would be hyperacute and have a uniform threshold, while judgements between modules would be limited to the coarse local-signs assigned to features. If this is true then one would expect that if the Vernier threshold is measured for various separation of the lines, the curve would be segmented into at least two regions, corresponding to within-module and between-module judgements. Westheimer (1984) reported data on precisely such an experiment using an interval discrimination task. He found that the the threshold rises smoothly from ~10″ at a 6′ interval to ~22″ at an interval of 12′. Although the data does not disprove Westheimer's hypothesis, it does cast some doubt on it. In fact, data (Levi & Waugh, 1996) shows that there are likely two separate regions, one where the gap is less than 24 arcmin, and the other where the gap is larger that 24 arcmin, as the plots tend to have a 'kink' around 24 arcmin and the slope changes.

2.1.1 Vernier Acuity and Orientation Discrimination

Westheimer (1981) discussed at length the relationship between Vernier acuity and orientation discrimination. In short, discriminating orientation and offset of two lines are merely two complementary descriptions of the same phenomena. This becomes clear when the effects of line length and separation on Vernier acuity are studied (Westheimer & McKee, 1977b). Increasing line length leads to lower thresholds, but as soon as the lines are longer than a few minutes of arc there is no more benefit. A similar result is reached when the separation between the lines is increased. At first there is no effect, but then as the separation grows beyond a few minutes the threshold increases. As might be suspected, two dots separated by ~3′ gives thresholds as low as any other Vernier stimulus. The threshold in the dot display is comparable to orientation discrimination thresholds for lines 3′ long. Westheimer (1981) explained that better orientation discrimination for longer lines is merely the result of hyperacute measurement of the displacements of the line ends. On the other hand, a filter model's perspective would argue that the longer lines also engage larger filters. With more filters engaged, thresholds should be lower from just probability summation.

Given the above considerations it appears that to detect Vernier offsets, a filter type model should try to detect non-vertical components for a vertical Vernier stimulus of orientation at the highest spatial frequency scale.

2.1.2 Vernier Acuity and Displacement Acuity

Hyperacuity can be found in other types of stimuli than Vernier acuity. If a single thin line is shown to an observer, and then the the line is suddenly shifted very slightly in the direction opposite to its orientation, then this shift can be detected at a hyperacute level. It is referred to as displacement acuity, and Badcock & Westheimer (1985b, 1985a) studied the effects of placing a flanking line in various positions around the target line.

The relationship between Vernier acuity and displacement acuity is not immediately obvious. In the Vernier acuity task, which can be described as a spatial misalignment, *orientationally adjacent* detectors at the *same location* are responsible for the low thresholds. In the displacement acuity task (temporal misalignment), it is the *spatially adjacent* detectors of the *same orientation* which detect the shift and are responsible for the low thresholds. The fact that detectors are perhaps more closely spaced in orientation than in space may explain why Vernier hyperacuity thresholds often are as low as 5-6″ whereas displacement acuity thresholds never go below 10-12″.

2.1.3 Spatial Frequency Results

The effects of isolated spatial frequencies on Vernier acuity has been measured using sinusoidal gratings (Bradley & Skottun, 1987; Whitaker & MacVeigh, 1991). The findings have been very much in line with the predictions of filter type models. Bradley & Skottun (1987) were not certain that some of their data at low spatial frequencies were compatible with filter type models, but the data in question were most likely due to an artifact of their experimental procedure.

2.1.4 Vernier Acuity for Modified Vernier Elements

A number of experiments have been performed in which the characteristics of the elements of the Vernier display have been modified (De Valois & De Valois, 1991; Kooi *et al.*, 1991; Toet & Koenderink, 1988). In all cases there was no effect on the Vernier acuity thresholds; it seems as the low spatial frequency characteristics of the envelope completely determined the threshold irrespective of the orientation (Kooi *et al.*, 1991), spatial frequency (Kooi *et al.*, 1991; Toet & Koenderink, 1988), color (Kooi *et al.*, 1991), or motion (De Valois & De Valois, 1991) of the element. The only interesting phenomenon was that stationary motion of the target induced a positional bias in the opposite direction.

All of the experiments mentioned use separations larger than the largest filter scale, which is about one degree of visual angle, and therefore filter theory is at a loss to explain these results. Perhaps some type of local sign is attached to these elements; in that case the local sign seems insensitive to many stimulus features. It is also clear that whatever process is making the offset judgement, it can operate across channels as well as within channels, although the process is somehow distance dependent.

2.1.5 Spatial Interference and Flanking Lines

Flanking lines were first introduced by Westheimer & Hauske (1975), and have been used to study interference in various hyperacuity tasks. Westheimer & McKee (1977b) found that flanking lines at about 2′ increased the Vernier threshold. More recent studies (Badcock & Westheimer, 1985a, 1985b) have shown that when the results of the experiments are analyzed carefully, it becomes apparent that there is no increase in threshold. There is, however, a shift in the bias as a result of the flanking lines and the effects of the flanking lines on this bias is much more interesting than the effects on the threshold.

Badcock & Westheimer showed that there is a nearby linear zone which attracts the target line if it is of the same contrast polarity as the flanking line, and repels the target if it is of the opposite contrast polarity. This linear zone extends only 3-4′ to each side of the target line. Beyond the linear zone is an inhibitory zone where the flank line always repels the target line irrespective of contrast polarity. Badcock & Westheimer (1985b) also claimed to show that the inhibitory zone extends far above and below a vertical target line. This result seems improbable however, and is probably due to the illusory completion of the two halves of the flanking line as they split the flanking line in two. In both the linear and the inhibitory zone one can place flanking line symmetrically around the target line without affecting the threshold.

The nearby linear zone fits very well with a filter type model, the inhibitory zone does not make any sense however, and it has no support in any version of the local-sign theory either. It *does* make sense in terms of the BCS model proposed by Grossberg & Mingolla (1985a, 1985b, 1987) and also in terms of the complex channel model (Sutter *et al.*, 1989; Graham *et al.*, 1992).

2.1.6 Other Types of Spatial Interference

In an interesting experiment designed to test the effects of introducing distractor dots on hyperacuity Morgan *et al.* (1990) produced some data which cannot easily be explained with a filter-type model. In a dot Vernier task with the dots separated by 10 to 40′, distractor dots of the same size were introduced between the target dots. The surprising finding was that the dots had little effect on the threshold unless they were extremely close to the target dots. A filter-type theory would expect that the distractor dots would interfere with the filters and thus produce higher thresholds, but this is not what happened. There was interference when the distractor dots were placed very close to the targets; this may indicate that filters of a very small size were interfered with. The inference is that small filters are used to localize the target dots, but the comparison process which produces the offset judgement is able to ignore the distractor dots. Morgan, Ward, & Hole (1990) argue that a hybrid explanation of the type mentioned is necessary to explain this data. Interestingly, this is also consistent with the BCS model of Grossberg & Mingolla (1985a, 1985b, 1987).

2.1.7 Adaptation Phenomena

It has previously been established (McKee & Westheimer, 1978) that the oblique effect has no significant influence on Vernier acuity. The oblique effect refers to the phenomenon that regular acuity is better in the horizontal and vertical directions than it is in oblique directions. McKee & Westheimer McKee & Westheimer (1978) studied the effects of practice on Vernier acuity and found that thresholds decreased at similar rates in all orientations tested (horizontal, vertical, and oblique). The differences between the thresholds in the different directions were always small, on the order of 3″. One of their subject showed a "vertical effect", vertical stimuli had higher thresholds than oblique stimuli. Only after two months of testing on vertical stimuli was the threshold in the vertical direction as low as the oblique threshold. Data from the near periphery (Westheimer, 2005) suggest that there exists an oblique effect for hyperacuity, thus the overall picture may be incomplete.

Watt & Campbell (1985) performed an interesting experiment with a Vernier acuity task. Instead of keeping the orientation of the stimulus horizontal or vertical at all times – the common way to perform Vernier acuity tests – they displayed the stimulus in a random orientation for each trial. The orientations were all within ±16° of the vertical, and the subjects had to report a rightward or a leftward shift. Subjects performance in this task was much worse than when the stimulus always appeared in the vertical orientation, assessed in a separate block of trials. This experiment is different from the one reported above in that the orientation of the stimuli are all very close; in the McKee & Westheimer experiment (1978), only three very different orientations were used (horizontal, vertical, and 43° oblique).

It is not clear what particular feature of the Watt & Campbell (1985) task caused the differential behaviour. The authors did not remark on this finding but used it to establish a baseline for other experiments. It appears as if the subject builds up some internal representation against which the presented stimulus is compared. When each new stimulus appears in a random orientation, the subject is prevented from constructing an accurate internal representation. It is unclear whether this built-up representation is a form of adaptation or some form of learning. In either case, little is known about the time-course of the phenomenon which also contributes to the difficulty of attaching it to a particular process.

A complete theory of pattern discrimination will have to account for effects such as the one demonstrated by Watt & Campbell (1985) either by including explanations in the theory or by including effects of a visual short-term memory (or whatever is responsible for these effects). At the present time, these effects have not been sufficiently well studied, and visual short-term memory is poorly understood (with many remaining unconvinced of its existence). Consequently it would difficult to propose a theory which can account for these effects. However, theories still need to be developed, and the important thing to remember is that the data reported in almost all experiments are for trained subjects; i.e., the data reported are optimal in the sense that training cannot further lower the thresholds. If this knowledge is applied to the filter type explanation of hyperacuity previously discussed, we can speculate that perhaps the subject learns to "pay attention" to a certain class of filters while ignoring all others.

Since the subject may not ever see the "correct" percept in some paradigms, one may wonder against what internal state the stimulus is compared. It is not unreasonable to suppose that the subject may build up an internal state which is the average of all recent stimulus presentations; this state may also be the standard state for making judgements about new stimuli. One problem with this approach is how to explain the first few judgements.

2.2 Models of Hyperacuity

As mentioned earlier, there are two main types of theories which have been put forth as explanations of hyperacuity. The first type will be referred to as local-sign theory, the second type will be called filter theory.

2.2.1 Hering

Local-sign refers to the localization of objects and features in visual space. In other words, the absolute position of features is somehow measured, and the feature is given a place-marker, or local sign. Later, local signs can be compared and thus displacements can be judged. This type of theory was accepted by psychophysicists at the end of the 19th century. When the cones in the retina were discovered, they were taken as the basic unit of distance in visual space. Hyperacuity therefore presented a difficult problem: if the cone spacing is the minimum unit of local sign how is it possible to make judgements with thresholds only a fraction of this distance? Hering (1899) was the first to suggest a solution to this dilemma. He proposed that local signs are averaged along the length of a line in Vernier display, and that is how fine localization judgements are made. This explanation can be generalized to include integration over time as well as space. Hering's explanation, with slight modifications, remained the state of the art until 1981. This is quite surprising because Ludvigh disproved Hering's hypothesis in 1953 by showing that Vernier thresholds using dots were as accurate as judgements with lines, as long as the dots are separated by a few arcmin (see section 2.1.1 on page 13).

2.2.2 Westheimer

Westheimer in his review of hyperacuity (1981) proposed a modified local-sign theory to explain this discrepancy. Westheimer was particularly impressed with data on interference (Westheimer & McKee, 1977b) which shows that adding irrelevant features to the stimulus can increase thresholds (i.e., subjects performance is worse). In particular, he attached special importance to the fact that the added irrelevant features must be placed fairly close to the target. His modified theory postulates that every feature receives a coarse local-sign, but there is a modular process which can produce very accurate relative localization judgements. This additional process is easily interfered with if there are too many features present in the region of visual space over which it operates. Westheimer (1981) also discussed the possibility of providing a fine scale local-sign for every feature, but discarded this option due to parsimony; also, the interference data argues for the modular approach.

2.2.3 Westheimer & Gelb: Line Element Theory

A very different type of theory was proposed by Wilson & Gelb (1984), and was applied to hyperacuity by Wilson (1986, 1991). Although several other authors have presented similar theories (Carlson & Klopfenstein, 1985; Daugman, 2007), the Wilson & Gelb theory is unique in that it based purely on human psychophysical data and has a small number of free parameters. Wilson and his coworkers performed masking studies to get spatial frequency tuning data (Wilson *et al.*, 1983) and orientation tuning data (Phillips & Wilson, 1984), that form the basis of the pattern discrimination model.

The basic insight of Wilson & Gelb was that detection and discrimination can have very different performances even though the same type of mechanism is used by both tasks. This is illustrated in Figure 2-1. The middle filter is very good at detecting stimuli in the middle of the stimulus range and poor at discriminating two nearby stimuli in the middle of the range. The opposite is true for the nearby filters; these are poor at *detecting* a stimulus in the middle, but show a large response differential to non-identical stimuli near the middle. This is exactly the mechanism which has been used to explain broadband color perception with only three cones.

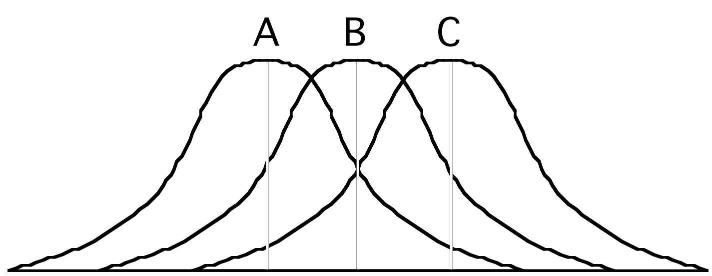


Figure 2-1: The difference between detection and discrimination in a line-element theory. The curves shown are response functions for three hypothetical filters along some physical dimension (horizontal axis). Filter B is ideal for detecting the presence of a stimulus in the middle. However, B is not well suited to discriminating the position of two stimuli near the middle since the response is about the same; this is easily accomplished by filters A and C because they show a large response differential.

This line-element theory explains many general features of hyperacuity data. For example, hyperacuity thresholds only seem possible when the stimuli are well above contrast detection threshold (Westheimer & McKee, 1977a); the theory explains this by noting that the filters which perform the discrimination only respond to stimuli which are well above threshold. Also, Westheimer & McKee (1977a) showed that Vernier acuity thresholds seemed insensitive to contrast above a certain threshold, this is also explained by the theory since the ratio of response of two neighboring filters is what used in the discrimination task. It must be realized that two-dot discrimination is not a discrimination task, as was mentioned before, it is really the detection of a luminance valley between two peaks formed by the sum of two point-spread functions. Since Vernier acuity is a discrimination task, that is why performance can be so much better.

The model of Wilson & Gelb (1984) was not intended as an explanation of hyperacuity only, but to explain other pattern discrimination data as well. The model can be summarized as follows: The first stage is a filter bank with a large number of even-symmetric receptive fields of six different spatial frequencies, ranging from 0.8 cpd to 16.0 cpd . These filters also cover 12 to 24 orientations, depending on the scale, and are closely spaced (every quarter cycle). All of these filters are convolved with the input. The signal of each filter is put through a compressive non-linearity in order to account for non-linear contrast behaviour. The outputs of every filter are compared with stored outputs of a “correct” input pattern. The Euclidean norm of this difference vector is the discriminability as computed by the model.

2.2.4 Klein & Levi, Carlson & Klopfenstein

The filters used by Wilson & Gelb have oriented receptive fields which are similar to Gabor functions but are constructed from separable Gaussian kernels. Klein & Levi (1985), for example, use a function described by Cauchy which is also similar to Gabor functions but has a frequency response which is symmetric when plotted in logarithmic coordinates. The model of Carlson & Klopfenstein (1985) is slightly different in that it is a signal detection type model which only analyzes the spatial frequency content of the stimulus. They do not assume a particular receptive field shape but assume a bank of filters which are fairly narrowly tuned to spatial frequencies. Both of these models are different from the Wilson & Gelb model in that they calculate the discriminability for *each channel* separately; the Wilson & Gelb model performs a vector difference within a *limited spatial region*.

2.2.5 Wilson

Wilson (1986) used the previously specified model (Wilson & Gelb, 1984) to explain a wide range of hyperacuity data. He nicely shows how the model predicts lower thresholds for increasing line lengths in the Vernier acuity task as demonstrated by Westheimer & McKee (1977b) among others. The model also predicts the increasing threshold with increasing separation of the lines in the Vernier acuity task. This curve (Westheimer & McKee, 1977b) is flat for separations smaller than about 6′ and rises for larger separations. The model predicts this curve very accurately. He also applies the model to the data on chevron acuity reported by Westheimer & McKee (1977b). Chevron acuity is very similar to Vernier acuity but different in that the top line is almost aligned with a chevron pointing towards the line. The model again produces a curve for threshold versus chevron angle which corresponds reasonably well with the data, although the thresholds in this case are a couple of arcseconds too high.

Wilson also applied the model to periodic acuity (Tyler, 1973), which is the discrimination of a straight line from a line with sinusoidal modulation. The model predicts the change in threshold with the spatial frequency of modulation very well. Measurements have been made of Vernier thresholds for sinusoidal gratings (Bradley & Freeman, 1985). Wilson computed the model's predicted thresholds as a function of spatial frequency and found the qualitative agreement to be good, the data showed that the actual thresholds were higher. When applied to data from the lab of Morgan, the model's predictions were right on the mark. The most likely explanation is that the measured thresholds of Bradley & Freeman (1985) were higher than they should have been because they used lower contrast gratings. Wilson also applied the model to data on interference of flanking lines on Vernier acuity (Westheimer & Hauske, 1975; Levi *et al.*, 1985). Again, the curve depicting Vernier threshold versus flank separation mirrors the data with an increase in threshold at a separation of 2′.

Finally, Wilson tackles the remarkable data of Klein & Levi (1985) where they showed that hyperacuity thresholds smaller than 1″ are possible. Klein & Levi employed a line bisection task with three parallel horizontal lines. On experimental trials the middle line was shifted slightly up or down, and the task of the subject was to identify whether it was closer to the top line, closer to the bottom line, or in the middle. The main manipulation in the experiment was the variation of the separation between the lines. With a separation of 1.3′ thresholds smaller than 1″ were obtained. Thresholds increase as the separation is decreased or increased from this point and there is a bump in the curve for separations around 2′. The model predicts the dip at 1.3′ and the bump at 2′ very well, and it also predicts further dips and bumps although none appear in the data. Wilson (1986) claims that the data is too sparse to rule out further dips and bumps, but the curve in the data looks very smooth at this point. The dips and bumps predicted by the model is a result of the different spatial scales; it is possible that the model of Wilson & Gelb puts too much emphasis on the discreteness of spatial scales. Klein & Levi have a different explanation for their data. When the lines of the display are as close as 1.3′, the size of the line-spread function starts to become very important. In fact, at this separation the three line-spread functions interact to produce a percept of three lines with the intervening regions sometimes becoming brighter than the background (the lines are bright on a dark background). If the top region looks brighter than the bottom region, then the middle line is shifted down. According to Klein & Levi (1985) this is really a brightness judgement, and not a localization judgement. Although it is not necessary for an explanation of hyperacuity to explain this data, it is reassuring to know that a filter bank based on psychophysical measurements is able to code these differences.

The precision with which the Wilson & Gelb model predicts the changes in thresholds for a variety of conditions is astonishing; especially in the light of the fact that a local-sign type model can make almost no predictions at all about most of these phenomena. The nice fit to such a large amount of data speaks for the usefulness of this type of model as an explanation of hyperacuity. It is important to remember though, that there are many problems with this model, both methodological and as applied to other data, which still need to be addressed.

2.2.6 Summary of the Wilson Model

The basic idea of the Wilson (1986) model, is that there are always a few filters which are well positioned so that small changes in input configuration produces significant changes in filter output for some of the filters. The filters used for detection of the stimulus are not the same as the filters used for judging the displacement of the stimulus, these latter cells are usually adjacent to the detecting cell as shown in Figure 2-1. While this scheme is an effective foundation for making hyperacute discriminations, more processing than just filters may be required when more complex stimuli are introduced.

2.2.7 Limitations of the Wilson Model

To produce a response the Wilson model relies on an internal representation of an "ideal" stimulus (in the case of Vernier acuity a straight line) which may never have been shown to the subject. The model does not specify how these internal stimuli are learned or otherwise represented innately. The idea of comparing a processed stimulus to an internal representation is not out of the question, but it depends on how this internal state comes about. In the case of Wilson's model, the modeler puts the ideal internal representation into the model, and the system computes a vector difference with this ideal.

In addition, the Wilson model only discriminates a difference from the internal "ideal" stimulus, it cannot distinguish between a 'right' or 'left' stimulus, which is crucial is some Vernier and other hyperacuity paradigms. A satisfactory model that explains hyperacuity *must* include such a mechanism.

The Wilson model, as it stands, is also inadequate for modeling data where there are spatial and oriented inhibitory interactions between individual elements of the stimulus (such as the data of (Badcock & Westheimer, 1985a, b; Morgan *et al.*, 1990; Westheimer & McKee, 1977b; Williams & Essock, 1986). Furthermore, there are data which require completion processes and feedback

(Toet & Koenderink, 1988), which are not approachable with the Wilson model. If the Wilson model were to be extended to include explanations for these phenomena, it would have to include some of the mechanisms which are part of the BCS, but it would still be unable to distinguish 'right' from 'left'.

2.3 Remaining Problems

A number of problems remain unexplained, here is a partial list:

1) There is still a need to explain some type of long-range comparison process. The data of Burbeck (1987) and Toet & Koenderink (1988) cannot be explained with any reasonable filter theory because the filters would have to be larger than the largest filters measured (Wilson *et al.*, 1983). It is clear that filters can be used to give each element of the stimulus a local sign, but there is still no understanding of the comparison process. The BCS model can solve some of these problems since it incorporates a long-range cooperative process.

2) The filter type model does not adequately explain the zones of inhibition observed by Badcock & Westheimer (1985a). These zones are non-linear and therefore do not fit in a linear filter theory. This is another case where a BCS type model could have some explanatory success. The inhibitory zones can be associated with the first competitive stage, which occurs after the rectification step. Furthermore, the data of Badcock & Westheimer (1985a) can easily be explained if the cooperative long-range feedback goes to cells in the first competitive stage, with subsequent inhibition.

3) The temporal integration data presented by Westheimer & McKee (1977a) is still a mystery. It is not clear how a BCS type model could explain these data.

4) The different properties of displacement threshold hyperacuity and Vernier hyperacuity need to be further explained. The fact that the displacement threshold is constant for all sizes and shapes of stimuli, even for different spatial frequencies should give a strong clue to the mechanism. The filter theory fails miserably in the displacement case even though it works well for Vernier acuity displays of different types, even for different spatial frequencies. Experimental paradigms such as reported by Hirsch & Mjolsness (1992) may shed light on this problem.

2.4 Experimental Directions

The database of hyperacuity experiments is large, and growing at a steady pace. Given the large amount of data, one would expect to find experiments which would provide a well defined domain for simulations with the hope of estimating the parameters of a model such as the BCS. This is not the case.

To perform such simulations, one needs data from experiments which are specifically tailored for this purpose. Specifically, one needs studies where carefully chosen stimulus variables are parametrically varied. Ideally, one would have one such study for each parameter of the model under study.

Furthermore, it would be ideal to have a paradigm where the parameters of the model could indeed be isolated for separate study. Understandably, it is difficult to find one experimental paradigm which satisfies these demands when studying a model as complex as the BCS. The paradigm used in this dissertation, the hyperacuity line displacement task, does provide some of the sought-after advantages.

As a consequence of the listed remaining problems, there are a few directions where there are opportunities for original work:

1) More data is needed on what has been called interference. The inhibitory zones of Badcock & Westheimer (1985a) need to be mapped out carefully, and should be possible using slight modifications of their displays. This research did not fully explore all the possibilities opened up by this technique. The published results (Badcock & Westheimer, 1985a, 1985b) also leave a number of questions unanswered. In particular, the extent of the inhibitory surround (both horizontally and vertically) of a vertical line was not sufficiently probed. It would be nice to establish experimentally that long-range completion is responsible for some of the unexpected data of Badcock & Westheimer (1985a). This is in fact what is reported in Chapter 3.

2) It would be interesting to construct band-pass filtered Vernier displays. In this way one could avoid stimulating the high spatial frequency filters which most often seem to mediate the perception of hyperacuity. It would be especially interesting to measure the extent of excitatory and inhibitory zones of such stimuli.

3) Another way to selectively disrupt the perception of hyperacuity is to use a masking paradigm. A specific filter selected for spatial frequency, orientation, and perhaps even spatial position can be masked and the effect on the hyperacuity theshold can be measured. This would be an invaluable tool for testing specifics of filter theories. This was done by Levi & Waugh (1996), and modeling work based on these results is presented in Chapter 5.

The technique for measuring distortions of the visual field due to a small number of elements is ideal for parametrically investigating various competitive, long-range completion, and feedback mechanisms which have been proposed as essential to the visual system (Grossberg & Mingolla, 1985a, 1985b). Lesher & Mingolla (1993) have found a qualitative role for these types of interactions when studying illusory contours.

Chapter 3: Hyperacuity Experiments

3.1 Introduction: Motivation of Experiments

A growing number of visual phenomena suggest that visual space does not have a rigid metric that is unaffected by the visual stimulus. Indeed, there appear to be some similarities with the effects of mass (i.e., contrast) on the geometry of space-time. Usually the effects are very small and only noticeable under special circumstances. In general relativity, it is sometimes possible to see light "bending" around massive objects. In the case of vision, the well-known Herring, Wundt, and Zöllner illusions all involve the bending of lines when presented in conjunction with other lines at an angle to the line. These illusions are also related to the mis-estimation of angles first studied by Jastrow (1891) and reviewed by Berliner & Berliner (1948). Other illusions such as the Café-wall illusion, Fraser's twisted-cord and spiral illusions fool the observer by creating perceived orientations which are not present in the stimulus. Although these examples are mostly concerned with intersecting lines, they demonstrate just how deformable the underlying geometry of visual space is.

Figure 3-1: The interval in A appears wider than the interval in B. The distance between the outer lines in both figures is identical, but the subdivided interval looks wider. After Illusion 4.2 in Resnikoff (1989).

Experiments on contours in after-images by Day (1962) and more recently on different types of contours by Rivest & Cavanagh (1996) also show that contours warp nearby visual space. In these experiments the deformation of visual space is evidenced by the repulsion or attraction of the apparent location of nearby contours. For example, Rivest & Cavanagh (1996) found attraction in all attribute combinations, but repulsion occurred only in some combinations, such as motion vs. luminance and texture vs. motion. Another effect which can be explained by the warping of visual space near contours is illustrated in Figure 1. The subdivided interval appears wider than the simple interval. If each line causes slight repulsion of neighboring lines, increasing the apparent separation between them slightly, then the combined effect of several lines can add up to become a noticeable increase in width.

A clever technique was pioneered by Badcock & Westheimer (1985a, 1985b). They used the line displacement hyperacuity task originated by Westheimer (1981) to measure these spatial interactions in the visual system. The stimulus has two time intervals, each containing a vertical line. The second line is shifted slightly to the right or to the left. The observer must respond either “left” or “right”. The threshold for this task is the amount of shift at which the observer can identify the direction of the shift 75% of the time. The line displacement task is a shift in time rather than a shift in space, as in Vernier acuity. The two paradigms are compared in Figure 3-2.

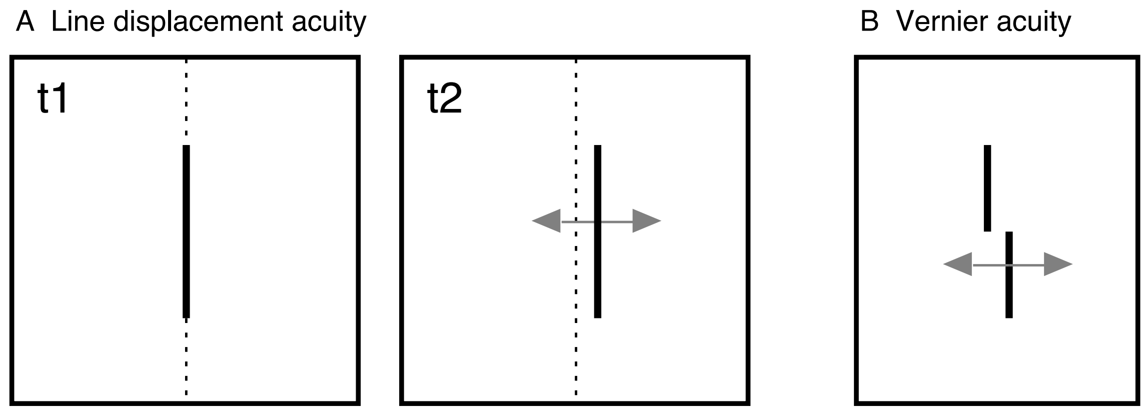


Figure 3-2: Comparing line displacement acuity and Vernier acuity. The displacement acuity stimulus in A has two intervals in time, one line is shown in each interval. The Vernier acuity stimulus shown in B also contains two intervals, but in space rather than time. The top half of the stimulus corresponds to the first interval of A, and the bottom half of B to the second interval in A. Note that neither the dashed line nor the arrows in the figures are part of the display; the dashed line is a reference so that the reader may infer the movement of the target; the arrows indicate responses allowed by the observer. Also note that the shift depicted is grossly exaggerated and that the stimuli are not drawn to scale, illustrating the spatial and temporal relations between figural elements.

The special feature of Badcock & Westheimer's (1985a, 1985b) experiments was that they introduced a flanking line of lower contrast during the presentation of the second (shifted) line. The flanking line introduced distortions into the visual field when placed near the target line. Badcock & Westheimer quantified the distortions by measuring the bias, i.e., the amount of shift required for the line *not* to move perceptually (see Figure 3-3), for each condition. For example if a flanking line is presented 6 minutes of arc to the right of the shifted target line, the two lines will repel each other and the target line will have to be shifted to the right about 10 seconds of arc in order stay in the same apparent position.

It is important to note that because the flanking line is presented only during the second time interval, any distortion of the apparent position of the line in the first interval is minimized. In contrast, when a flanking line is introduced in the Vernier task, it distorts the perceived position of both the upper and lower lines as shown in Figure 3-4. Given the difficulty of separating the effects of the flanking line in the Vernier case, especially when the flanking line is moved up or down, only the displacement paradigm is used in this study.

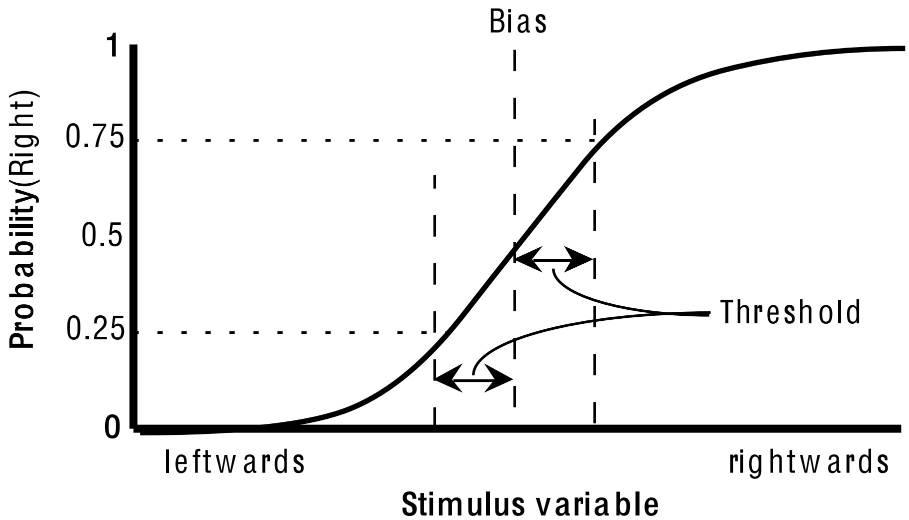


Figure 3-3: Bias and thresholds are here defined in terms of performance levels. The subject responds “left” or “right” to each stimulus. If the stimuli corresponding to the 25% and 75% performance levels can be established, the bias is the midpoint and the threshold is half the distance between the two stimuli.

The flanking line might possibly affect the perceived position of the target line in the first interval through iconic storage or backwards masking. However, after a presentation of 500 ms, the position of the target line is quite stable; which is also why the target is presented alone in the first interval and the flanking line introduced in the second interval. Furthermore, if there is any effect of the flanking line on the target in the first interval, then the effect will always be to slightly underestimate the bias. This underestimate would not be expected to vary much across conditions in this paradigm.

The line displacement paradigm is hyperacute, meaning that it allows detection of differences smaller than than the center-to-center distance between cones in the fovea (~25 seconds of arc). It is a general technique that can be used to study many types of interactions in early vision. The experiments reported here are direct extensions of previous work (Badcock & Westheimer, 1985a, 1985b), examining more extensively the horizontal and vertical extents of the inhibitory zone. New experiments are also reported, testing the effects of asymmetrical flanks and of an illusory flanking line. Together these experiments are designed to constrain models of early vision (Grossberg & Mingolla, 1985a, 1985b, 1987; Peterhans *et al.*, 1986; Finkel & Edelman, 1989), for an overview see Lesher (1995). For example, special attention has been paid to finding parametric measures of the extent of the inhibitory zone around line-detectors, both horizontally and vertically.

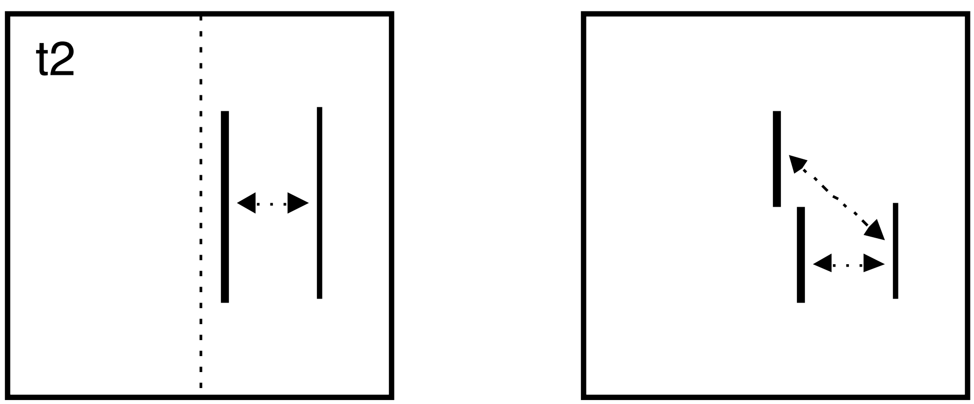


Figure 3-4: The main advantage of the displacement paradigm is that the flanking line does not have an opportunity to distort the perceived position of the line present in the first interval, since they are not present simultaneously. In the Vernier paradigm (right), it is difficult to distinguish the effects of the flanking line on the upper and lower Vernier lines. In this figure, dashed lines with arrows denote effects of attraction or repulsion.

3.2 Methods

The basic structure of the stimuli is the same as the displacement acuity task used by Badcock & Westheimer (1985a, 1985b) described above. The stimulus consists of two intervals in time. A thin vertical target line, (high contrast on a dark background) is displayed near the middle of the monitor. After 500 ms the target line shifts to the right or to the left, and this line is displayed for another 500 ms. The task of the observer is simply to specify whether the shift was to the left or to the right. Responses where the response times were longer than 2500 ms were discarded. When the target line is shifted, other structures may be introduced which may alter the perceived position of the target line. Badcock & Westheimer (1985a, 1985b) used a vertical flanking line of lower contrast displaced from the target line. Clearly a variety of other carefully chosen configurations can be chosen instead of a simple flanking line. Several configurations are used in the experiments reported here; the particular configuration used for each experiment is shown and discussed in the corresponding section below.

Display: The screen on the SGI Indigo is 1280 x 1024 pixels, and the width and the height of the display image were adjusted so that pixels were square, 0.27 mm on a side. Observers were seated so that the distance from their eyes to the screen was 500 cm. Thus pixels were 11.1 arcsec square. The size of the entire display was almost 4 deg wide by 3 deg high, which provided ample viewing area to find the edges of the inhibitory zones.

The thin lines were 0.5 arcmin thick, meaning that on the average a line was three pixels wide. Anti-aliasing was used so that the lines could be positioned arbitrarily on the screen (256 positions within each pixel); thus a thin line generally consisted of two partial and two full pixels. The luminance was linearized by gamma lookup and the maximum brightness was 65 cd/m2 and a black screen measured 0.03 cd/m2.

Data Collection: A total of at least 200 trials were collected for each condition. In two instances 250 trials were collected: observer YA in Experiment 2, and the parametric conditions of Experiment 4. All the trials collected for each condition were used to estimate the psychometric function. Each session consisted of 350–450 trials and lasted just under half an hour. Each trial was initiated at the response of a previous trial, and the stimulus was preceded by a small dot for 1500 ms showing the location of the target line in the central area of the display. No feedback was provided in any experiment due to the nature of the experiments. The observers preferred to do the sessions in pairs; thus the experiments required between four and five hours of observation over a period of about a week.

A new combination of previously existing data collection methods was used. Because it was important to find both the point of subjective equality (PSE) – when the number of 'right' responses equals the number of 'left' responses – as well as the discriminability threshold (75%), the whole psychometric function had to be estimated for each condition. The Adaptive Probit Estimation (APE) (Watt & Andrews, 1981) method was used for selecting stimulus levels, and Maximum Likelihood Estimation (MLE) was used to fit the psychometric function as in Best PEST (Pentland, 1980) and QUEST (Watson & Pelli, 1983). The logistic function was used because of its analytical simplicity as well as being the form of the psychometric function resulting from two different theoretical analyses (Luce, 1959; Link, 1992), and also because it is indistinguishable from the cumulative Gaussian in practice. Standard errors of the estimates derived using MLE were computed using the Bootstrap method described by Foster & Bischof (1991).

The APE method performs preliminary estimates of the psychometric function and uses these estimates to decide what values to use for the next set of trials (10). Each trial is randomly chosen from four possible values, two on each side of the midpoint of the preliminary psychometric function. Two points are near the middle (thus more difficult for the observer) and two are farther away from the midpoint (easier for the observer). The points are randomly chosen from this set, but the method also ensures balanced coverage. While the method efficiently homes in on the best range of values to use for a given subject for a given condition, it is difficult to display the collected data and its fit to a psychometric function, because data is collected for so many different points with only a small number of trials for each stimulus value. This problem and one suggested remedy, binning, are illustrated in Figure 3-5 below. The top plot shows the raw data with one point for every stimuls level used; the clearer bottom plot shows shows the same data collected into equally spaced bins centered on the best-fitting function.

The MLE procedure merely finds the most probable logistic function from which all the data collected for a condition could have come. To calculate the error estimate, the Bootstrap method uses the psychometric function estimated from the data to generate one hundred simulated data sets. These data sets have the same number of measurements at the same stimulus levels as the collected data set, but with data generated by a pseudo-random function and the best-fitting psychometric function. Each of the simulated data sets are used to estimate a separate psychometric function. Then statistics are computed for the collection of simulated functions to provide error estimates for the estimated function. This method works because the statistics of the simulated psychometric functions are assumed to be closely related to the true statistics around the estimated function. This assumption is reasonable as there is no expectation that the distribution of errors would change appreciably for related psychometric functions (Press *et al.*, 1986).

Further details of the new Adaptive Logistics Estimation (ALE) method can be found in the section on Psychophysical Methods in Appendix C (p. 151).

Figures: In the figures, what is plotted is the measured shift, i.e., the bias point of each condition. Badcock & Westheimer (1985a, 1985b) labelled their figures 'Induced Shift', but what is really being measured here is a real shift, the shift in the bias of the target line to null out the effect of the flanking line.

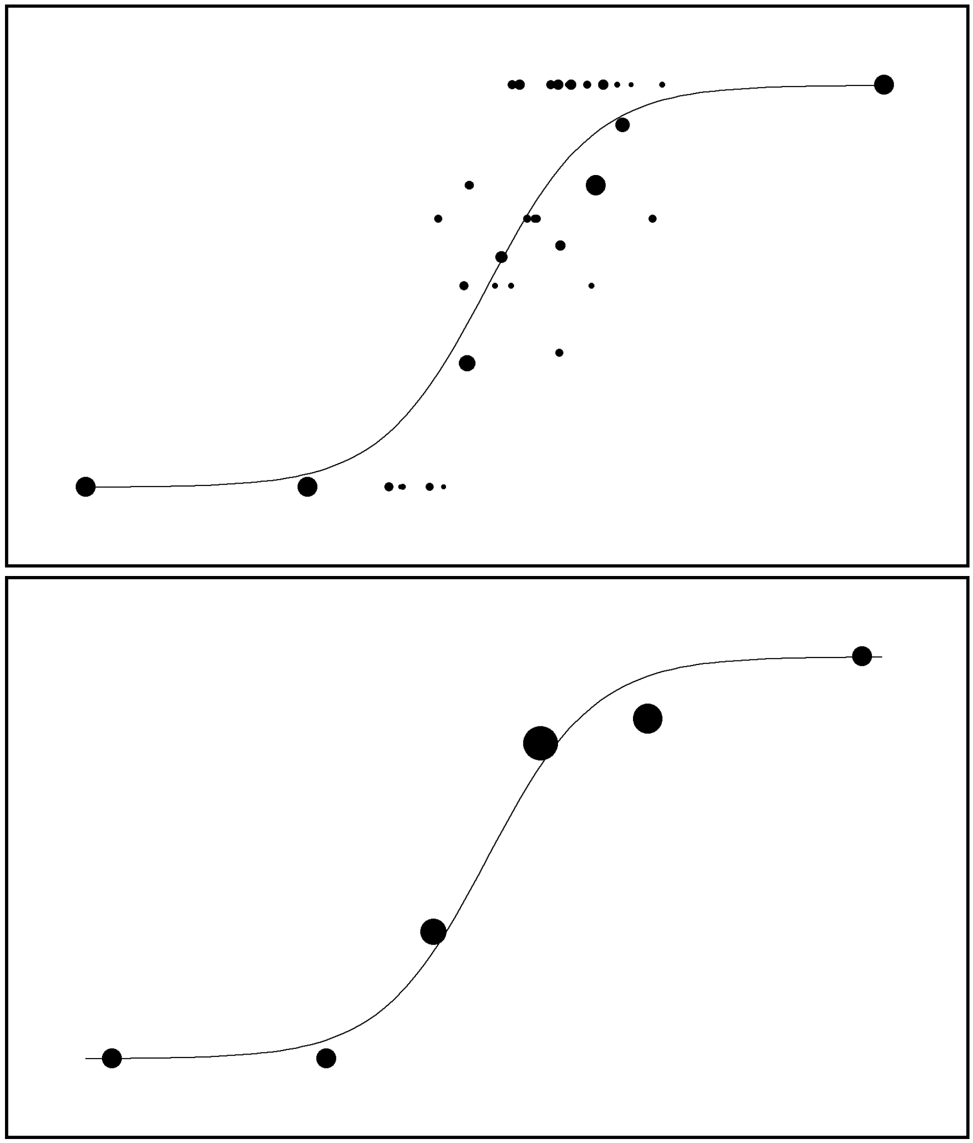


Figure 3-5: Two ways to plot the data collected using the APE method. The area of the dots correspond to the number of trials collected at that level. The top figure shows the raw data with a point for every stimulus level used. The bottom figure shows the same data collected into equally spaced "bins" centered on the best-fitting function. Notice that it is much easier to visually estimate the "goodness of fit" in the bottom figure.

3.3 Experiment 1

This experiment establishes the baseline performance in the condition where there is no flanking line. It is also a control condition and a verification of the method. It measures the direction of shift threshold as well as any inherent bias. This bias may be due to processing in the visual system or perhaps due to higher level decision making processes. Figure 3-6 below shows the stimulus configuration for this experiment. The target line is shifted either to the right or to the left and no flanking line is added.

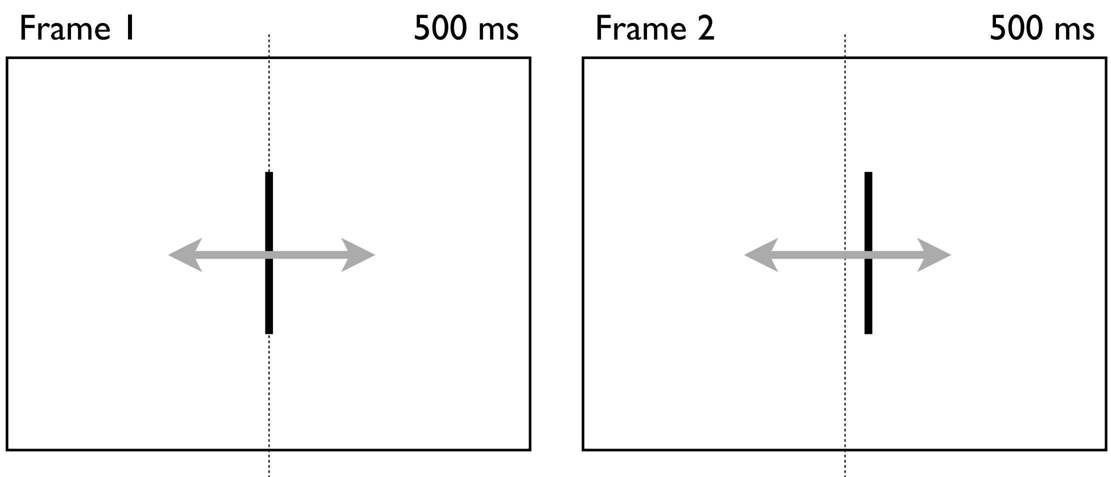


Figure 3-6: The two frames depict the first 500 ms and the second 500 ms of the stimulus. Neither the dashed line nor the arrows are part of the display; the dashed line is a reference so that the reader may easily see the movement of the target; the arrows indicate allowed responses by the subject. The displacement in the figure (~1 arcmin) is exaggerated because the thresholds are so small. Actual experimental displacements are much smaller. The same exaggeration is present in all figures in this chapter showing a displacement.

The results of this experiment are shown in Table 3-1 below. The threshold for the most experienced observer (HR) is at the lower range of the values reported by Badcock & Westheimer (1985a, 1985b). The other observer (DJ) has a threshold which is higher than those of experienced observers. This is expected because perceptual learning is an important effect in these types of tasks (Westheimer, 1981; Saarinen & Levi, 1995). Only very experienced observers have thresholds in the 10″ range. In this experiment and those that follow, only one observer (HR) can be considered experienced, the others have not had the same amount of practice In summary, the thresholds compare well with previously reported thresholds.

Table 3-1: Results for Experiment 1.

Observer	No Flank
HR	5.5 ± 1.62 (9.0)
DJ	–2.6 ± 3.00 (18.4)
Legend: (used for all tables)	PSE" ± standard error (threshold")

The inherent bias measured is significant for one observer (HR). While the experimental conditions were randomized and there was no expectation of bias, it could easily be the result of visual processing or decision making processes. There is no expectation that any bias would disappear during training, as there was no feedback during either practice or experimental trials. In any case, the bias is small enough not to interfere with results of the experiments that follow. The slight positive bias is also visible in Figures 3-8 and 3-14.

For modeling purposes this experiment could be used as an estimate of the noise which needs to be added to the outputs of the model.

3.4 Experiment 2

To map out the size and shape of the inhibitory zone discussed by Badcock & Westheimer (1985a), a single flanking line is used (see Figure 3-7 below). Badcock & Westheimer found that reversing the contrast polarity of the flanking line did not affect the induced shift in the inhibitory zone. Therefore, because only the inhibitory zone is of interest, there is no need to reverse the contrast polarity. In this experiment, a flanking line with half the contrast of the target line and also half as tall, (i.e., 5 arcmin) was added to the second frame of the stimulus (again, see Figure 3-7 below). The horizontal separation varied from 5 to 50 arcmin and three vertical locations were used: vertically centered, just above the target line with a zero tip-to-tip vertical separation, and a tip-to-tip vertical separation of 5 arcmin. See Table 3-2 below for representative diagrams of the stimulus configurations.

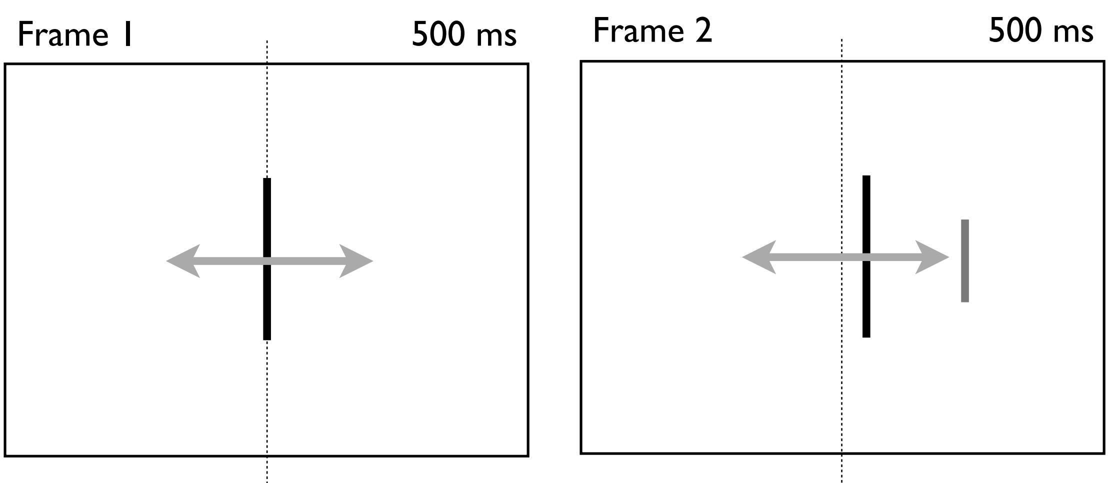


Figure 3-7: The symbols have the same meaning as in Figure 3-6 above. In this experiment a flanking line (the shorter gray line) is introduced during the second time interval. The experimental conditions consist of varying the horizontal and vertical position of the flanking line.

The results of the experiment are shown in Figure 3-8 and the numerical values are listed in Table 3-2 below. Even though the flanking line is only half as tall as the flanking line used by Badcock & Westheimer (1985a), the amount of repulsion is the same (ca. 10 arcsec). The strongest trend in the data is the decrease in repulsion as the separation between the flank and the target is increased. This is apparent both horizontally and vertically. In terms of the horizontal separation, there is a maximum repulsion at 5′ or 15′ (depending on the condition) and then a tapering off to (in most cases) a minimal effect at 50′. The repulsion is at a maximum when the flanking line is vertically centered and completely 'overlaps' the target line, but again tapers off with increasing separation. If one takes into account the inherent bias measured in Experiment 1, then it is clear that the plots for one observer (HR) shows a lot of symmetry. The plot for the other observer (YA) is markedly less symmetrical, most likely due to not being as experienced with this hyperacuity task.

These results show that the inhibitory zone around the target line extends farther than the maximum separation (8.4′) used by Badcock & Westheimer (1985a). In fact, the *peak* of the repulsion in many cases occurs at a separation of 15′ and does not disappear until the separation is 45-50′.

In terms of significance for the modeling effort, this experiment maps out the size and shape of the spatial extent of the inhibitory zone of the competitive stage.

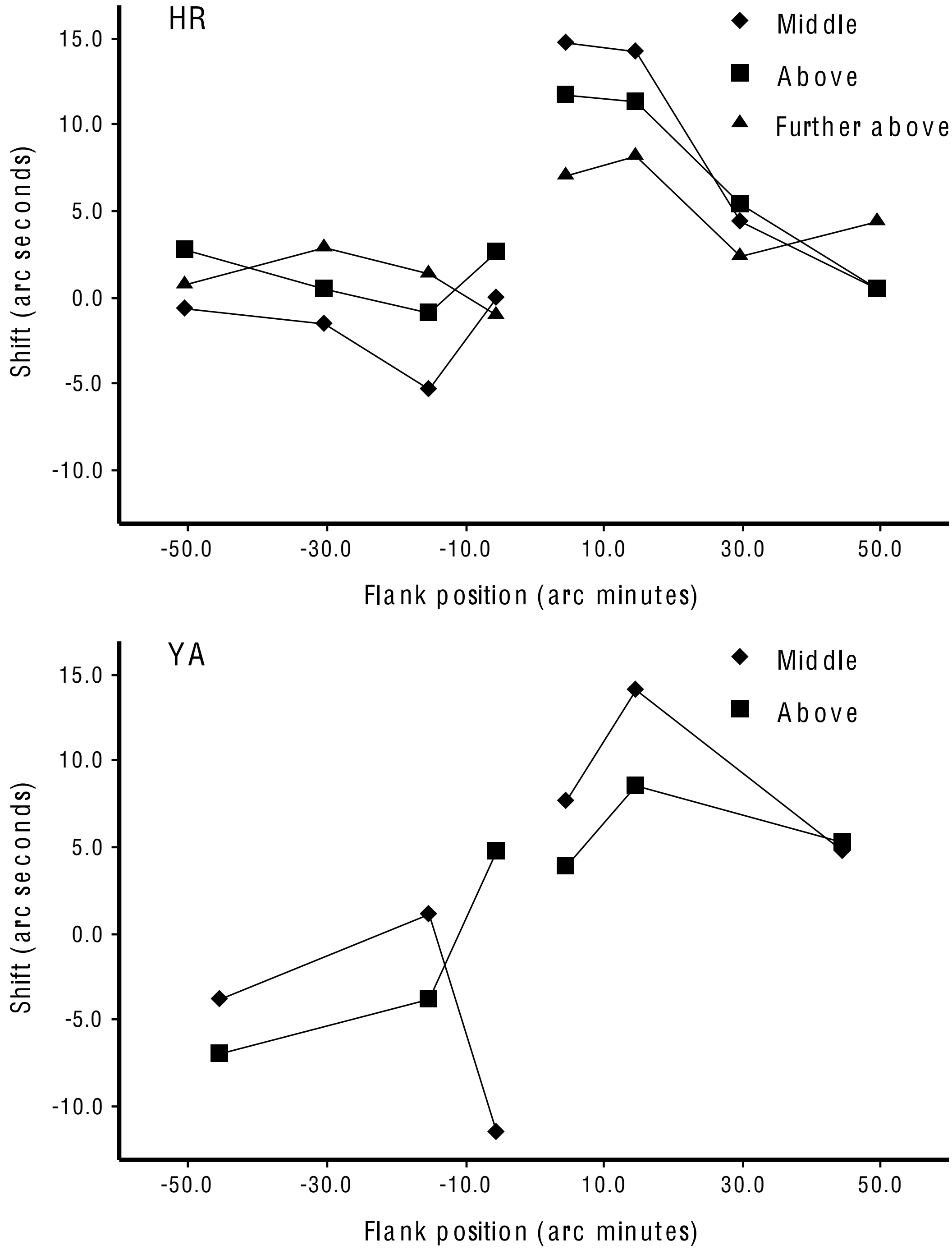


Figure 3-8: Plots of the shifts required for the conditions of Experiment 2. The middle sections are not connected as this covers the zone of attraction, which has very different properties and was not tested.

Table 3-2: Results for Experiment 2 (See Table 3-1 for legend.)

Observer	Flank position (arcmin)							
HR	-50	-30	-15	-5	5	15	30	50
	–0.7 ±1.27 (7.8)	–1.6 ±2.05 (13.2)	–5.4 ±1.80 (11.7)	–0.1 ±1.52 (8.9)	14.7 ±1.50 (8.9)	14.2 ±1.47 (8.8)	4.3 ±1.87 (10.6)	0.5 ±1.34 (7.5)
	2.7 ±1.67 (8.9)	0.5 ±1.66 (9.1)	–1.0 ±1.94 (10.0)	2.6 ±1.67 (10.8)	11.7 ±1.30 (7.2)	11.3 ±1.91 (10.5)	5.4 ±1.60 (9.3)	0.5 ±1.59 (9.5)
	0.7 ±1.20 (7.2)	2.8 ±1.84 (11.2)	1.3 ±1.97 (11.4)	–1.1 ±1.69 (10.9)	7.0 ±1.65 (10.2)	8.1 ±1.35 (8.8)	2.3 ±1.51 (9.4)	4.3 ±1.15 (7.2)
YA	–45		–15	–5	5	15		45
	–3.8 ±2.39 (15.3)		1.1 ±2.56 (19.1)	–11.6 ±3.13 (19.2)	7.6 ±3.87 (27.2)	14.1 ±2.52 (17.8)		4.7 ±2.95 (18.4)
	–7.0 ±2.51 (15.9)		–3.9 ±3.14 (18.1)	4.7 ±3.01 (18.0)	3.8 ±3.38 (23.2)	8.5 ±3.28 (21.3)		5.2 ±3.29 (21.8)

Note — Fewer conditions were used for observer YA to reduce the total time.

3.5 Experiment 3

Another experiment was designed to test the 'splitting' of the flanking line (Badcock & Westheimer, 1985a) to see if there is some kind of 'linking' or long-range com pletion between the two collinear fragments. This type of linking, described by Schumann (1900, 1987), would be akin to the gestalt perception of a dashed line as one long line.

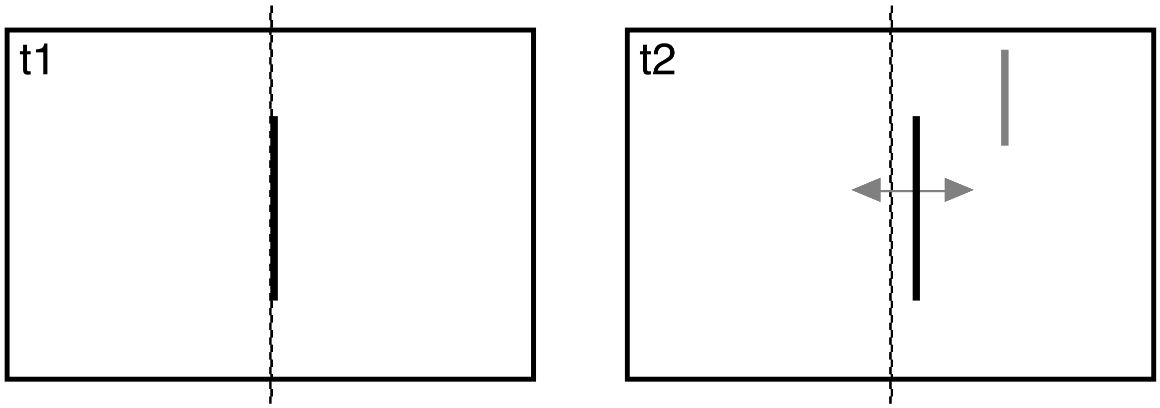


Figure 3-9: Test the effect of half a flanking line located above or below the target. The prediction is that there is no (or a very small) inhibitory effect. In this way, the vertical extent of the first competitive stage can be measured. The displacement has been exaggerated for clarity.

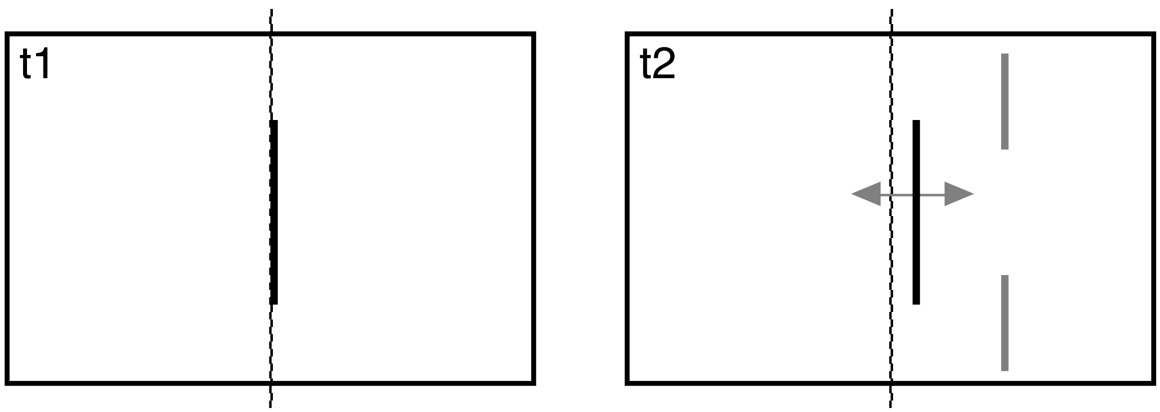


Figure 3-10: In the second condition, two collinear flanking line halves are used. The prediction here is that there will be an inhibitory effect much greater than twice the effect above; and that it is due to long-range completion. The displacement has been exaggerated for clarity.

In one set of conditions (Figure 3-9), a single half of the flanking line was used, and in the other set (Figure 3-10) both halves were used; representative diagrams are also included in Figure 3-11 and in Table 3-3. Badcock & Westheimer (1985a) tested the latter set of conditions, which are symmetrical, either with the two fragments joined into one flanking line or with the two flanking line fragments separated vertically. The new conditions are those using only one half of the flanking line, and are thus vertically asymmetric. The working hypothesis for the present experiment was that a single half-flank would have no effect on the target location when located above or below the target line. The assumption behind this hypothesis is that there is a completion process that takes place before any spatial inhibition. Thus with a single half flank there is no linking contour in the inhibitory surround zone, and there should be no repulsion.

The results for this experiment are shown in Figure 3-11 and are listed in Table 3-3 below. The data for the symmetrical conditions support those reported by Badcock & Westheimer (1985a) and show clearly the zones of attraction and repulsion. When the full flanking line is very close to the target (1.2′ or 1.0′) there is attraction, and when the flanking line is a little farther from the target line (6.0′ or 5.0′) there is repulsion, which is also consistent with results from the previous experiment. When the same target-flank separations are tried with the flanking line split into two collinear halves with a tip-to-tip vertical separation equal to the length of the target line (10′), there is always repulsion.

In the asymmetrical conditions, with only the upper or the lower of the two flank-halves, there is also always repulsion. Furthermore, while the pattern for observer HR shows clearly that one of the two halves of the flanking line causes the same amount of repulsion as both halves together, the data for observer DJ also points in this direction.

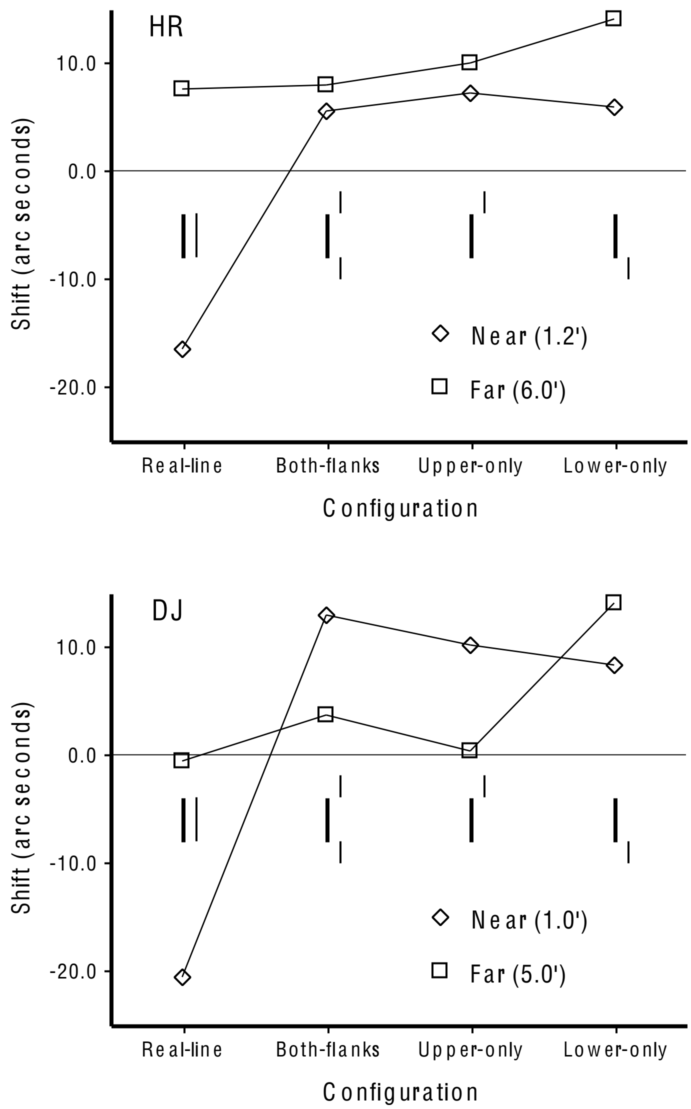


Figure 3-11: Results for Experiment 3. The attraction and repulsion for an overlapping flanking line ('Real-line') depends on the distance to the target. In the other conditions, it apparently makes no difference whether one or both flanking line halves are used, the amount of repulsion is the same. In this figure, the data from presentations on the left and the right have been combined according to Figure 1 in Badcock & Westheimer (1985a).

Table 3-3: Results for Experiment 3 (See Table 3-1 for legend.)

	Flank position (arcmin)							
Obs:	HR				DJ			
Cond.	–6.0	–1.2	1.2	6.0	–5.0	–1.0	1.0	5.0
	–2.1	15.3	–17.7	12.8	3.3	17.9	–23.3	2.0
	±1.98	±1.91	±1.85	±1.38	±2.85	±3.14	±4.44	±3.11
	(12.0)	(11.0)	(11.3)	(8.5)	(16.4)	(21.1)	(24.5)	(17.4)
	–3.1	–3.2	7.7	12.7	–7.0	–22.6	3.3	0.4
	±1.48	±2.48	±1.31	±1.69	2.78	±3.86	±4.23	±2.78
	(8.3)	(13.2)	(7.3)	(9.3)	(15.0)	(23.1)	(25.9)	(16.2)
	–9.2	–4.2	10.0	10.6	–3.1	–16.7	3.5	–2.4
	±2.51	±2.06	±1.67	±1.51	2.38	±2.52	±3.57	±3.00
	(14.8)	(12.5)	(10.9)	(8.3)	(14.2)	(13.8)	(21.8)	(17.9)
	–11.5	–8.1	3.5	13.2	–17.3	–9.7	6.6	10.7
	±1.90	±2.91	±1.91	±1.45	3.50	±4.39	±5.00	±2.91
	(12.0)	(15.2)	(10.8)	(7.6)	(20.9)	(22.3)	(28.2)	(15.7)

3.6 Experiment 4

While the luminance-defined flanking lines used in the previous experiments have been shown to cause repulsion and attraction of the target line, an interesting question is raised by the existence of ‘illusory lines’. Illusory lines are found in such phenomena as the boundary of a Kanisza triangle, the central part of an Ehrenstein figure and the ‘abutting grating illusion’ (Soriano *et al.*, 1996), as in Figure 3-12. In these phenomena, the illusory line is not part of the stimulus but is easily recognized as a contour by the visual system. The illusory lines referred to here are actually illusory edges with zero thickness. They often appear extremely sharp, sometimes sharper than real edges.

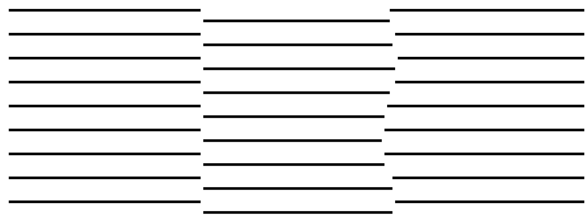


Figure 3-12: Examples of the abutting grating illusion. Notice that the left and right edges of the figure also give rise to illusory lines, albeit less salient than those in the middle.

This experiment was designed to test whether such 'illusory' lines also cause attraction and repulsion when presented next to a target line as in the previous experiments. The abutting or offset grating stimulus was chosen as the stimulus to provide the illusory line because of its salience and because there is no brightness effect associated with the illusory line. To produce the desired vertical illusory line, a horizontal grating is required. Moreover, to avoid biasing the induced shift of the target line when the grating is introduced, and also to avoid providing any clues about the direction of the shift, the lines making up the grating extend all the way to the edges of the display.

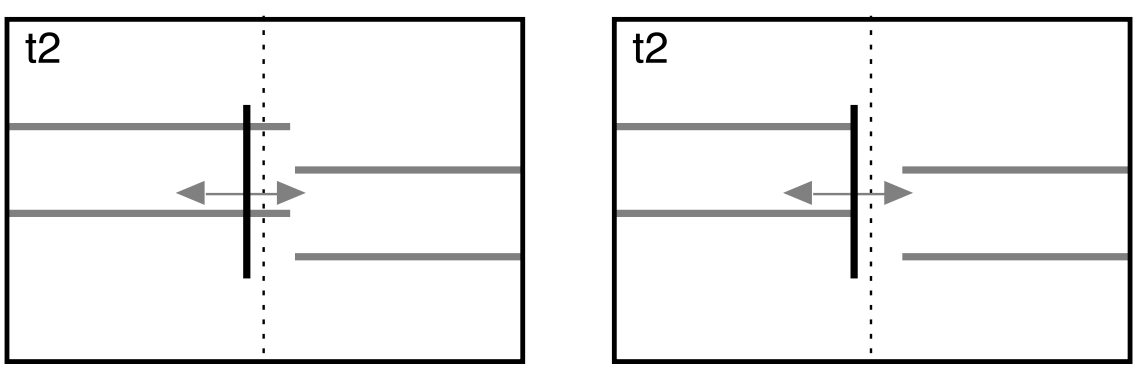


Figure 3-13: Stimuli used in Experiment 4. The line condition, similar to the internal contours of Figure 3-12, is on the left side. The gap condition, comparable to the outside edges of Figure 3-12, is on the right. In these figures as well, the displacement has been exaggerated for clarity.

This means that the lines making up the grating will pass through the target line and thereby slightly increase the mean luminance of the background. The appropriate controls for this stimulus must therefore be conditions where the offset in the grating occurs right on top of the target line (position 0.0) or entirely outside the display area (position $\pm\infty$).

The experiment uses two conditions, the 'gap' and 'line' variations shown in Figure 3-13. The line condition uses an abutting grating stimulus as the flanking line; the gap condition uses only one side of the grating. In both cases the separation between the target line and the flank is varied as in Experiment 2. The grating lines are 0.5 arcmin wide and of the same intensity as the flanking lines in Experiments 2 and 3. The line condition is most important, as it is the pure grating. The gap condition includes the effect of less luminance on the side of the gap as a result of the single-sided grating inducer.

The results are shown in Figure 3-14 and are also listed in Table 3-4 below. Figure 11 shows the results of the parametric variation of the gap and line conditions. This graph has several interesting features.

The line condition has a uniform light distribution along the horizontal axis (except of course, for the target line) and the shifts are due to the location of the illusory line alone. There is clearly repulsion between the illusory line and the target line at a separation of 15 arcmin. There are also indications of attraction between the illusory line and the target when the separation is very small, at 1 arcmin, although this effect appears to be asymmetric. When the illusory line is not present in the display, there is neither repulsion nor attraction.

The curve for the gap condition has a shape that is similar to the shape of the curve for the line condition. It tends to follow the line condition curve, but

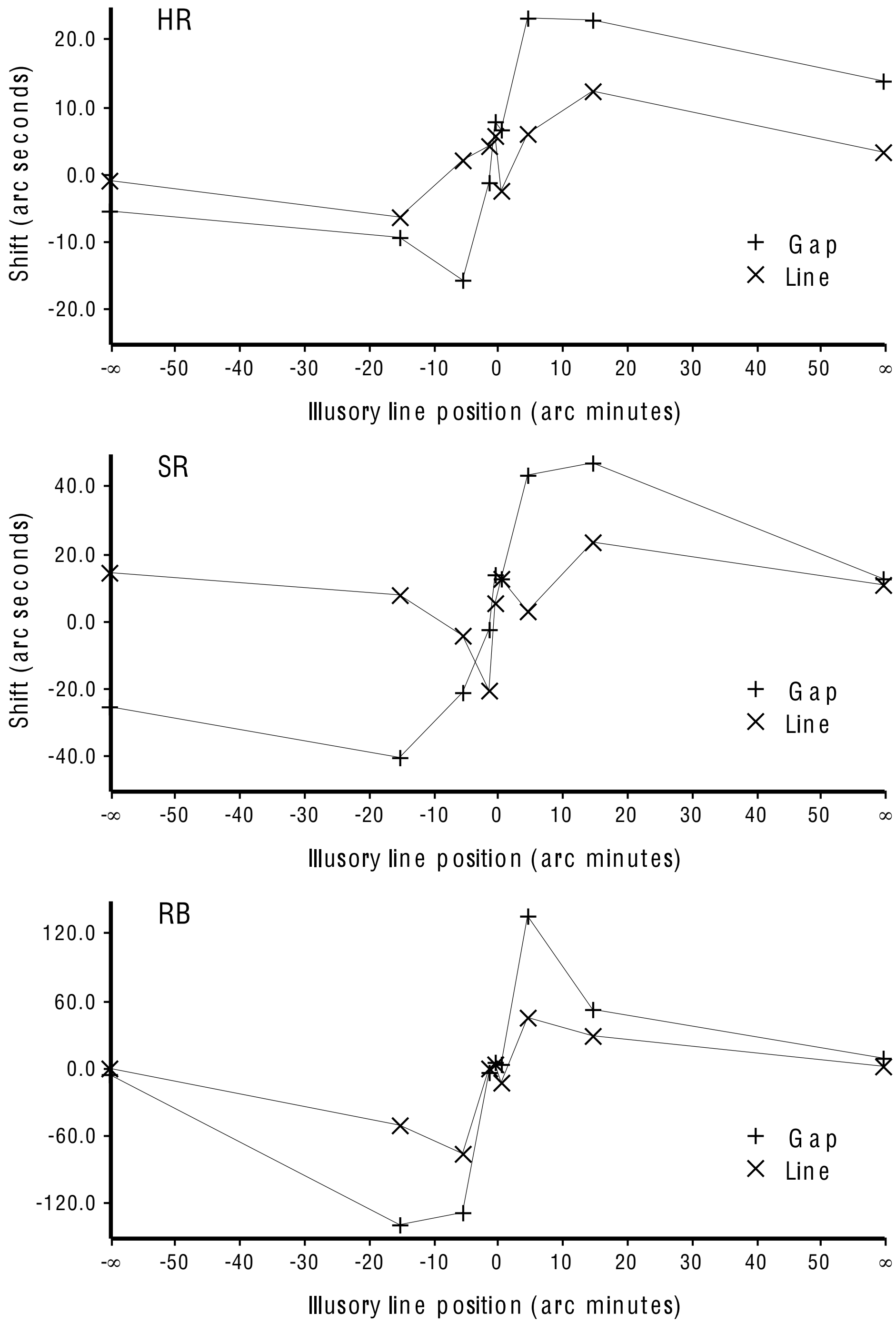


Figure 3-14: Results for Experiment 4. The effect of repulsion is readily apparent for all observers. The gap condition tends to exaggerate the repulsion observed in the line condition.

Table 3-4: Results for Experiment 4 (See Table 3-1 for legend.)

	'Illusory line' position (arcmin)								
	–∞	–15	–5	–1	0.0	1	5	15	∞
HR: Line	1.0 ±1.56 (8.3)	–6.3 ±1.76 (10.7)	1.9 ±1.79 (12.4)	4.0 ±2.53 (12.6)	5.6 ±1.70 (8.8)	–2.6 ±2.10 (12.3)	5.8 ±1.98 (10.4)	12.3 ±1.78 (10.0)	3.1 ±1.69 (10.2)
HR: Gap	–5.6 ±1.78 (9.5)	–9.3 ±2.03 (12.0)	–15.9 ±1.70 (10.3)	–1.3 ±2.15 (12.0)	7.9 ±2.39 (13.4)	6.7 ±1.37 (9.4)	23.1 ±1.99 (12.4)	22.8 ±1.78 (11.2)	13.8 ±2.32 (12.5)
SR: Line	14.4 ±5.10 (31.4)	7.7 ±5.25 (29.0)	-4.5 ±5.17 (31.5)	20.5 ±5.30 (27.7)	5.5 ±5.72 (33.7)	12.4 ±4.34 (24.9)	2.8 ±4.47 (25.1)	23.6 ±5.22 (29.0)	10.8 ±3.97 (22.5)
SR: Gap	-25.4 ±6.45 (37.2)	-40.9 ±3.98 (24.4)	-21.1 ±3.58 (21.0)	2.8 ±5.25 (28.5)	13.5 ±3.90 (23.5)	12.2 ±4.04 (22.4)	43.3 ±5.50 (30.8)	47.1 ±3.99 (22.6)	12.5 ±4.09 (23.3)
RB: Line	0.5 ±2.74 (15.7)	-51.1 ±10.61 (59.6)	-77.0 ±6.69 (39.4)	0.3 ±3.70 (20.2)	2.5 ±2.57 (14.6)	-12.9 ±5.63 (33.2)	45.2 ±5.43 (29.9)	28.1 ±5.13 (28.2)	1.5 ±2.71 (15.5)
RB: Gap	-6.6 ±3.18 (17.7)	-140.0 ±18.99 (112.8)	-129.5 ±13.80 (87.3)	-3.4 ±3.75 (22.3)	4.4 ±3.25 (19.1)	3.0 ±4.69 (27.9)	135.8 ±16.31 (91.7)	52.7 ±4.54 (24.6)	8.5 ±3.67 (21.0)

it shows even greater repulsion at separations in the range 5-15′. This difference is probably explained by the asymmetric distribution of light and dark in the central linear zone (Badcock & Westheimer, 1985b).

In summary, it is apparent from both the line and the gap conditions that there is repulsion between the illusory contour and the target line. There are also weaker indications that there is attraction between the illusory contour and the target line when in the central zone.

3.7 Discussion

Experiment 1 and other experimental conditions which are equivalent to (Badcock & Westheimer, 1985a, 1985b) conditions compare well with previous data. Hyperacuity thresholds for trained observers are close to previously published values, and the pattern of bias measurements also correlates well. The hyperacuity thresholds obtained also correlate with data from the rest of the literature. Inexperienced observers have larger thresholds but show patterns of bias that are consistent with those of experienced observers. Thresholds are fairly constant across conditions, with one exception. One observer in Experiment 4 showed some large thresholds in conditions where the bias was also large; this again, is most likely due to the inexperience of this observer.

Possible explanations for the types of phenomena that have been reported in this paper point to the use of an initial stage of local spatial frequency selective filters, and the subsequent combination of these filter outputs. In particular, the combination of filters must provide for lateral interactions between filters. Lateral inhibitory interactions are necessary in order to explain the inhibitory zone around perceived lines. In addition, it is likely that there is some type of linking between collinear filters which are spatially separated but which could be part of the same contour (Field *et al.*, 1993; Levi & Waugh, 1996).

The parametric data from Experiment 2 can be used to estimate the horizontal and vertical extents of the inhibitory surround. The data from one observer (HR) appears to be more symmetric than that of the other observer. When this data is used, and the shape of the inhibitory surround is approximated by a two-dimensional Gaussian, then the width of the Gaussian is 17.0 arcmin and the height is 11.5 arcmin. These numbers are only representative of the stimuli very similar to the target line. The filters most sensitive to the target line are vertically oriented filters with a spatial frequency in the range of 10-16 cyc deg-1.

However, it is not clear from the data reported here that the inhibitory surround is necessarily symmetric. There are indications of asymmetry for all observers, although it appears to be stronger for observers with less experience. Perhaps the inhibitory zone becomes more symmetrical with training. It is possible that this is another perceptual learning effect similar to the well known decrease in threshold with practice for Vernier and related tasks.

In any case, the field of inhibition is wider than it is tall, but only by a factor of about 1.5. On the other hand, the field is very wide compared to the size of filters at the smallest scale; a vertically oriented filter is only about 4 arcmin wide and usually somewhat taller.

It is very interesting to compare these results with those of Yu & Essock (1996), who used the very different 'Westheimer paradigm'. They measured the detection threshold of a line placed on a rectangle while varying the shape of this rectangle. They found 'strong end-zone antagonism beyond the ends of the elongated summation area, as well as flank antagonism to the sides.' There was threshold elevation (desensitization) for light within a central region, 5-6′ wide by 10-11′ tall, and threshold depression (sensitization) for light outside this region but within an end-tapered 14′ by 23′ region. These regions correspond well with the excitatory and inhibitory zones found Badcock & Westheimer (1985a), the size of the central zone matches very well. There is close agreement about the vertical extent of the inhibitory zone, but Yu & Essock found a much more limited horizontal extent than what was found in Experiment 2. One possible explanation for this might be that in their paradigm, the amount of inhibition saturates when the rectangular region reaches a certain size; thus further increases in width may not have an effect if the threshold was already maximally sensitized.

Lesher & Mingolla (1993), following a suggestion by Grossberg (1987), have found a qualitative role for the inhibitory interactions observed between parallel lines. They studied the strength of illusory figures generated from collinear line-ends, and found that the strength of the illusory contour is an inverted-U function of the line spacing. They explained their results using the concept of inhibitory zones, and their findings are entirely consistent with the results reported here. Recent modeling efforts successfully demonstrate this inverted-U function as a function of line spacing (Grossberg *et al.*, 1997).

Experiment 3 was designed as an explicit test of a prediction that the author had made based on the Boundary Contour System (BCS) theory (Grossberg & Mingolla, 1985a, 1985b). The expected explanation was that the two collinear line fragments would link to generate an invisible contour between them. The generated contour would be in the inhibitory surround of the target line and thus the two lines would repel each other. Removing one of the line fragments ought to destroy support for the invisible contour, and there should be no repulsion.

No reduction in the amount of repulsion was observed. This is not what was expected from the hypothesis presented above. Since there was no reduction in the amount of repulsion when one of the flanking lines was removed, the experiment gives no insight into whether the collinear fragments do link together and what effect this link might have on the positions of nearby lines. I.e., the link may exist but not have an additional effect on the target line. Levi & Waugh (1996) have proposed a such linking mechanism based on data from Vernier acuity experiments.

In the case of Experiments 2 and 3, the observed repulsion of the target line qualitatively requires no further mechanisms than the inhibitory surround described above. It is a little surprising, however, that a flanking line above (or below) and a little to one side of a target line should cause the target line to be repelled in the other direction almost as much as if the flanking line was just to the side of the target line. This could be the result of a saturation effect.

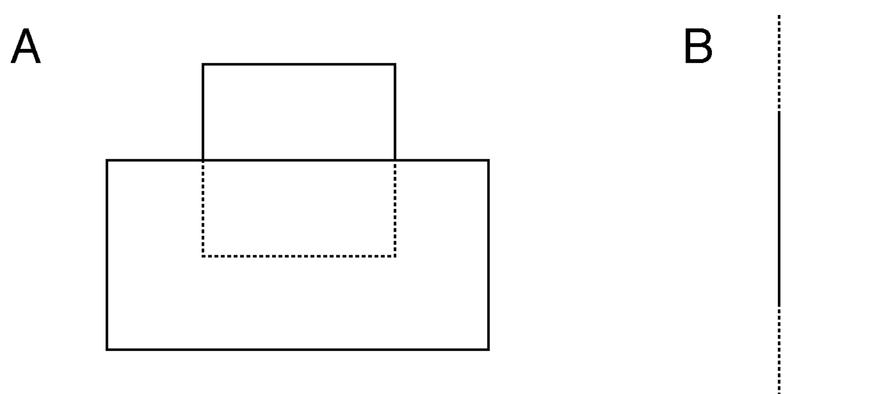


Figure 3-15: Possible amodal completion explanation of the results of Experiment 3. The small abutting rectangle in A has an amodal presence that has been investigated by Nakayama *et al.* (1989) If the ends of thin lines have a similar amodally completed continuation as shown in B, then this hypothesis can explain the observed repulsion.

An altogether different explanation of these phenomena is that the abrupt end of a thin line causes the same type of amodal completion seen when a small shape shares a border with a larger shape (see Figure 3-15). This latter type of amodal completion is quite common and has been investigated by Nakayama *et al.* (1989). The possibility of a line end giving rise to the same type of phenomenon, as if the rest of the line were hidden under an edge of the same color as the background, is not completely far-fetched, as it would also help explain the large vertical extents of the inhibitory zone.

The author proposes that the ends of thin lines induce amodal completion and the amodally completed lines always cause repulsion, i.e., an invisible component of a line continues past the visible end, as if the line were inserted into an imaginary slit in the background. This hypothesis would account for the results of Experiment 3 where a single half of the split flanking line caused as much repulsion as both halves together. The proposed hypothesis would also account for the surprisingly large vertical extent of the inhibitory zone.

The results from Experiment 4 are easier to explain. The main result from the experiment is that a flanking line formed by the abutting grating illusion has the same repelling effect on a target line as a high contrast luminance-defined flanking line. The behavior in the inhibitory zone is very similar, but the behavior in the central zone – where a luminance-defined line shows strong attraction – is less clear. The data suggest that the illusory line exhibits only a slight amount of attraction or repulsion, with no clear pattern for observers or conditions. Because illusory contours and real lines have the ability to cause repulsion of a target line when placed in the inhibitory zone, illusory contours and real lines may be more similar than previously thought.

Models of pre-attentive vision that explain the formation of lines from abutting gratings or other similar stimuli (Grossberg & Mingolla, 1985a, 1985b, 1987; Peterhans *et al.*, 1986; Finkel & Edelman, 1989) already predict or can be modified to incorporate the repulsion observed between the illusory flanking line and the target line.

Chapter 4:
Establishing BCS Parameters

The goal of this chapter is to reduce the number of unknown parameters used in the description of the BCS system. Since ultimately these parameters will need to be estimated by fitting the output of the model to observed psychophysical data, it is a good idea to eliminate or otherwise fix as many of these parameters as possible. The task of fitting data will become substantially more practical as the number of free parameters is reduced. Two separate methods will be used to reduce the number of free parameters, dimensional analysis, and single-stage input/output analysis using simple simulations.

4.1 Methods for Establishing BCS Parameters

Dimensional analysis of the equations provides a methodical approach to eliminating unnecessary parameters that have no effect on the dynamics of the system. As will be seen, the eliminated parameters usually perform some type of redundant scaling function.

In addition, it is also possible to analyze some of the stages of the system separately. For each stage there are certain input-pattern/output-pattern relations which are desirable. By adjusting the parameters over wide ranges it is possible to figure out the relationships between the parameters and the desired input and outputs. For some parameters this will limit the acceptable ranges of the parameter, this is clearly very useful for establishing the values of these parameters. In some cases, the valid range might even be small enough that the parameter is effectively a constant. In other cases, there may be no limit to the parameter at all, leaving it to be determined by other means.

Another goal of this section is to isolate those parameters which may have some basis in psychophysics or physiology from those parameters which are specific to the simulation of the system. In particular, the “discretization” of nodes in terms of the number of orientations and the spatial spacing are considered simulation parameters. That is, even though values of these parameters which are in some sense consistent with psychophysics and/or physiology are used, they are still determined by the simulation. One is tempted to describe these parameters as being “hardwired” into the simulation. In any case, it is precisely because those parameters are unavoidably fixed within the simulation that the description of the system being simulated should separate them from other parameters which are more changeable.

This is not commonly done. For example, a model may simulate 12 orientations, and specify a Gaussian orientation kernel with a σ of 2. That type of specification is not directly transferable to other simulations which do not use 12 orientations. Also, it has very little meaning in terms of psychophysical or physiological measurements which normally are specified in terms of degrees (°). It makes more sense to say that the σ is 30° and give it an explicit meaning that may carry outside of the simulation.

4.2 Dimensional Analysis

All of the equations in the model have a very similar form, thus we take a basic equation and apply the dimensional analysis procedure to it. Then we can generalize the results to the whole system. Therefore we start with a standard shunting equation

$$\frac{d\tilde{x}}{d\tilde{t}} = -\tilde{D}\tilde{x} + (\tilde{U} - \tilde{x})(\tilde{C}\tilde{\mathbf{i}}' * \mathbf{c} + \tilde{E}\tilde{f}) - (\tilde{x} + \tilde{L})\tilde{S}\tilde{\mathbf{i}}' * \mathbf{s} \ , \tag{4-1}$$

where $\tilde{x}$ is a variable quantity that is changing over time, $\tilde{D}$ is the time-constant of decay, $\tilde{U}$ and $\tilde{L}$ are the upper and lower limits of $\tilde{x}$, $\tilde{C}$ and $\tilde{S}$ are the scale factors for the center ($\mathbf{c}$) and surround ($\mathbf{s}$) kernels of the input ($\mathbf{i}'$), and $\tilde{E}$ is the scale factor for feedback inputs $\tilde{f}$. After dimensional analysis this becomes simplified to

$$\frac{dx}{dt} = -x + (1 - x)(C\mathbf{i}' * \mathbf{c} + f) - (x + L)\mathbf{i}' * \mathbf{s} \ , \tag{4-2}$$

where most of the parameters (uppercase) have vanished and

$$C = \tilde{C}/\tilde{S} \text{ and } L = \tilde{L}/\tilde{U}. \tag{4-3}$$

Remarkably, in this particular case it appears that six parameters have been reduced to two! This is not exactly true; as will be explained below, as two of the parameters have actually migrated to the equations of other stages. Several observations and comments are warranted at this point.

First it is necessary to establish what happened to the four parameters that disappeared; why are they not necessary? Two parameters, the $\tilde{D}$ and the $\tilde{U}$, establish scales for two variables. The $\tilde{D}$ is the one parameter in Equation 4-1 which sets the time-scale, and since the time-scale is arbitrary, it might as well be unity. Thus $\tilde{D}$ can be removed because is a completely arbitrary parameter which can be set to anything, but changing it will not help us understand the system. If at some point one would like to directly associate the mechanisms described by these equations to actual neural processes, then the $\tilde{D}$ would of course have a very explicit meaning in terms of the capacitance and therefore also the time-scale of neurons, but that is a matter that can be ignored here.

The $\tilde{U}$ parameter similarly establishes a scale for the values of the variable $\tilde{x}$. Again, it is arbitrary, so why not choose unity. As above, it would only have meaning if the system was considered in the context of being isomorphic with some neural processes, and $\tilde{x}$ would then represent the cell voltage or perhaps the cell spiking rate, and $\tilde{U}$ would then be an upper bound.

It is assumed by the procedure of dimensional analysis that the inputs to the equation ($\mathbf{i}'$ and $\tilde{f}$) can have all possible values, therefore it is assumed that they do not need to be scaled. This causes the $\tilde{E}$ to disappear, as well as either one of $\tilde{C}$ or $\tilde{S}$ ($\tilde{S}$ was chosen here). Most inputs clearly do not have this property (unlimited range), especially since they are the output of another equation with an upper bound of unity. However, every signal is transformed by a signal function, a useful and generic one seems to be

$$x' = M[x - \tau]^+ \ , \tag{4-4}$$

where x is the input, τ is a threshold, M is a scaling parameter, and the function $[\,]^+$ returns zero for all negative inputs and returns the input itself otherwise. In most cases, the threshold (τ) is set to zero. Each signal can be scaled arbitrarily by the M parameter however, thus satisfying the implicit assumption above. Note that this does add another parameter (M) for each variable.

We also have to consider what happens when several of these equations are put together into one system which incorporates feedback. The first concern is the time-scales of the equations; they could potentially all be different. However, in the BCS, there is no reason to assume that there is any difference in time-scales. The equations should all have the same time-scale.[5]

Thus at least two parameters are eliminated from each equation. For each variable there seems to be three parameters left: the ratio of the lower bound of activity to the upper bound (L); the ratio of the center to the surround (C); and of course the scaling of the variable as it enters other equations (M). These are all dimensionless variables.

In some cases it may be necessary to use recurrent equations instead of the purely feedforward equation (4-1). The recurrent equation has inputs from its own layer via the center and surround kernels as well as inputs from the previous layer and has the form

$$\frac{d\tilde{x}}{d\tilde{t}} = -\tilde{D}\tilde{x} + (\tilde{U} - \tilde{x})(\tilde{C}\tilde{\mathbf{x}}' * \mathbf{c} + \tilde{E}\tilde{f} + Gi') - (\tilde{x} + \tilde{L})\tilde{S}\tilde{\mathbf{x}}' * \mathbf{s} \ , \tag{4-5}$$

where the variables and parameters are the same as in Equation 4-1 ($\tilde{\mathbf{x}}'$ taking the place of $\mathbf{i}'$), with the addition of the G scaling factor for the input (i') from the previous layer. After dimensional analysis becomes

$$\frac{dx}{dt} = -x + (1 - x)(C\mathbf{x}' * \mathbf{c} + f + i') - (x + L)\mathbf{x}' * \mathbf{s} \ , \tag{4-6}$$

where, just like before, most of the parameters disappear and again,

$$C = \tilde{C}/\tilde{S} \text{ and } L = \tilde{L}/\tilde{U}. \tag{4-7}$$

There seems to be no difference in the number of parameters used in either case, but since the output of this stage (x') is also used as an input, the output scaling parameter (M) affects both this and the following stage.

---

5. There would be an exception if transmitter dynamics were employed in the dipole stage. The transmitters would be governed by equations at a slower time-scale than the node activations. The benefits of including these modifications are explored by Francis (1993).

4.3 Analysis of the Oriented Competition

The competitive stages of the BCS can now be analyzed individually in order to further reduce the number of parameters that need to be fitted. The oriented competition (most often referred to as the second competitive stage) is the simplest of the stages in the CC–loop, and is therefore a good place to start.

The oriented competition has two functions: (1) to sharpen any orientation signals present in the system, and (2) to help to produce so called “end-cuts” at the ends of lines. The mechanism is that spatial competition from nearby nodes cause inhibition below the tonic activity level in a node, and then the oriented competition enhances the signal in the opposite orientation. The equation that define this stage is identical to Equation 4-2. The center and surround kernels have a Gaussian shapes defined by

$$c = N_c \exp(-\left(\frac{k - k_0}{\sigma_c}\right)^2) \text{ and } s = N_s \exp(-\left(\frac{k - k_0}{\sigma_s}\right)^2) \,, \tag{4-8}$$

where N_c and N_s are scaling parameters, k is the orientation of the input node, k_0 is the orientation of the filter[6], and σ_c and σ_s are the widths of the Gaussian kernels. The k, k_0, σ_c and σ_s are all measured in the same unit of angle, preferably degrees. Note that orientations use modular arithmetic and wrap around at 180°; in particular, the difference between two orientations should always be in the range [–90°, 90°]. Equation 4-2 can be rewritten as a steady-state solution by solving for the variable x when there is no change. This gives

---

6. While the k_0, σ_c and σ_s may look like they have indices, they do not. The subscripts denote orientation of the filter, center kernel, and surround kernel, respectively. The next chapter introduces and explains a consistent notation.

$$x = \frac{C\mathbf{i}' * \mathbf{c} + f - L\mathbf{i}' * \mathbf{s}}{1 + C\mathbf{i}' * \mathbf{c} + f + \mathbf{i}' * \mathbf{s}}\,, \tag{4-9}$$

with the same parameters as described above.

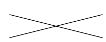
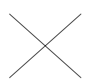
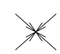
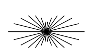
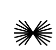
The basic operation of this stage is to impose a center–surround competition in orientation space. Thus the center has a narrow orientation bandwidth, and the surround has a wider bandwidth. In order to maximize the the end-cuts generated by the stage the center kernel should be close to zero at 90°, whereas the surround kernel should be as large as possible at 90°. These conditions are fulfilled when the center is relatively narrow and the surround is flat. Therefore, the surround kernel width can be assumed to be very large, leaving us with one less parameter to worry about.

Another desirable property of the oriented competition is that if all the orientations have equal activation (as when only the tonic input activates the first competitive stage), there should be no output at all. This property is also called featural noise suppression (Grossberg, 1983). By looking at Equation 4-9 it is obvious that this property can only obtain when $L \geq C$. This is yet another theoretical restriction on the parameter choices.

It is time to consider what the input/output relationships of this stage should be. It suffices to analyze the activation pattern of all orientations at one single spatial location. Consider six different input patterns listed in Table 4-1 (all have tonic activity at all orientations as well, of course), and the desired output activations of each input pattern. Activation of a node refers to the above-threshold activity only.

What follows are a number of simulations using these inputs and showing the corresponding outputs for various parameter sets. The parameters that have been varied are: the width of the excitatory kernel (σ_c, denoted by *s* in the plots); the input scaling factor (M); the ratio of the excitatory kernel to the inhibitory kernel (C); the ratio of the lower bound to the upper bound (L); and the tonic activity (T).

Table 4-1: List of desirable Input/Output relationships for the oriented competition (second competitive stage). Figures are included to demonstrate the orientation distributions.

#	Input	Desired Output
1	A single orientation active	Strong activation of that one orientation, and some activation of neighboring orientations 
2	Two lines at an acute angle (say about 30°) 	Strong activation in a group of neighboring orientations 
3	Two lines at a right angle 	It would be desirable if both of the orthogonal orientations would have some activation, and no activation in other orientations 
4	A broad pattern of inputs biased towards one orientation 	Output activations should be sharpened so that only a few orientations have any activity with the same bias as in the input pattern
5	Tonic activity in all orientations except no activity in a couple of neighboring orientations (the node is receiving inhibition from neighboring spatial nodes) 	Activation in orientations perpendicular to those with no tonic input, i.e., end–cuts 
6	Identical activity in all orientations 	No activation in any orientation

The first plot (Figure 4-1) shows the effect of different widths of the excitatory kernel (σ_c). When σ_c is small the output pattern is simply a scaled version of the input pattern. When σ_c is larger, it works more like a smoothing operator, spreading activity from one input orientation to neighboring output activations. For σ_c approaching 90°, activation is spread so uniformly that it becomes a uniform pattern which is then normalized away. A good choice of σ_c seems to be 15–20°; a value of 18° have been chosen for the remaining simulations.[7]

Figure 4-2 shows the effect of varying the other parameters. For each parameter (M, C, and L), higher and lower choices of the value are shown in an attempt to establish the range of behaviour. If M is too small, then there is very little output activity; for very large M, the transformation becomes a pure ratio scale. A reasonable choice for M seems to be 50, which is close to a pure ratio scale.

As mentioned above, from considerations of featural noise suppression it is necessary that $L \geq C$. In the simulations shown in Figure 4-2 it is apparent that output activations are severely diminished when $L > C$ to the point that end–cuts (5th column of patterns) disappear. On the other hand, outputs seem enhanced when $L < C$, such that end-cuts start to look very good; however, the uniform input is not suppressed as it should be. Changing the yoked values of C and L together has an effect almost identical to that of M. The outputs are normalized, becoming a ratio scale for larger values. The best choice for C and L is around 5.

7. This value also corresponds closely to a width at half–height of 25° for the kernel, a value very similar to the physiologically measured orientation tuning of visual cortical cells.

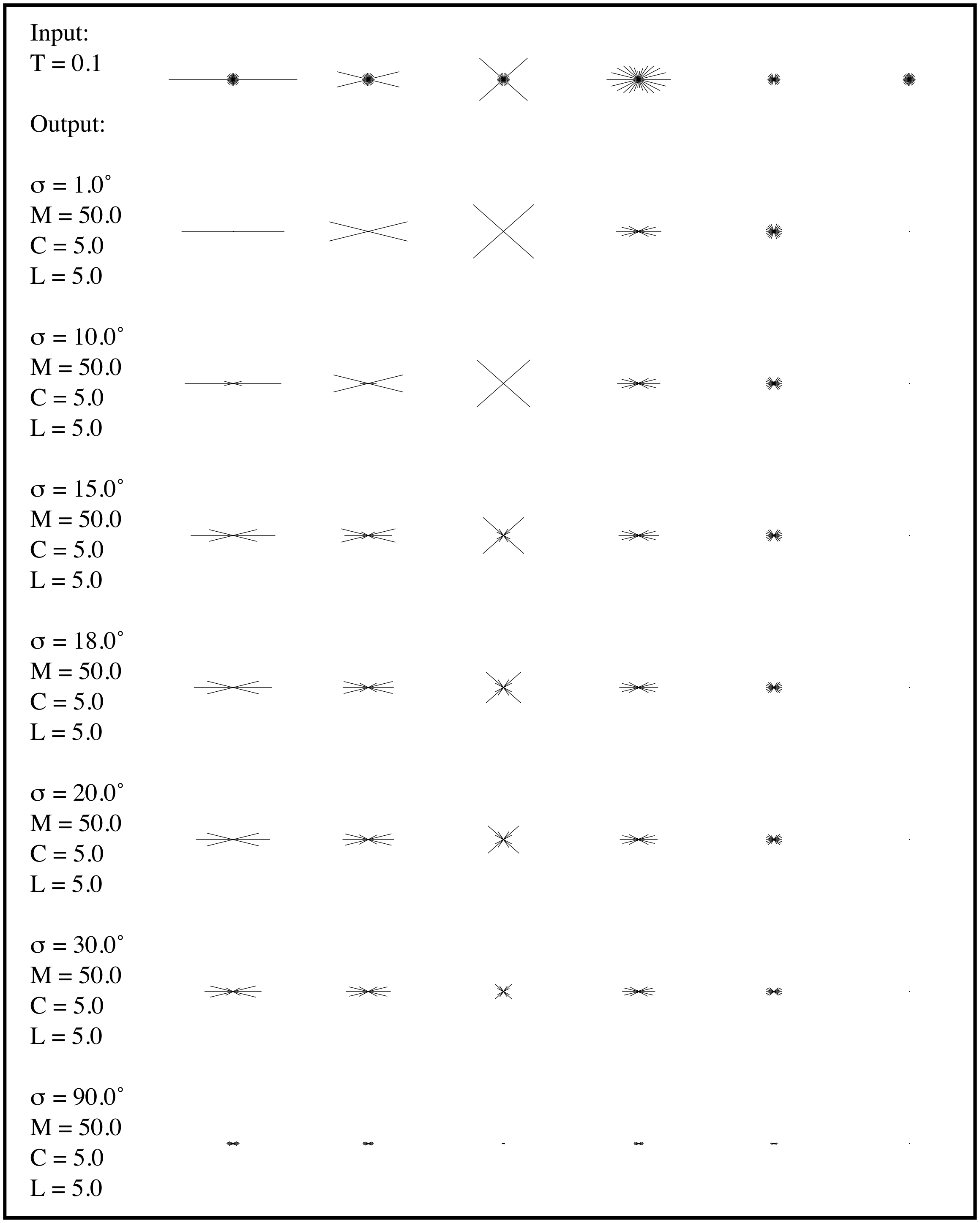


Figure 4-1: The output of the oriented competition given six different input patterns. Seven different widths of the excitatory kernel (σ_c) are used. It seems that values in the range 15–20° work best.

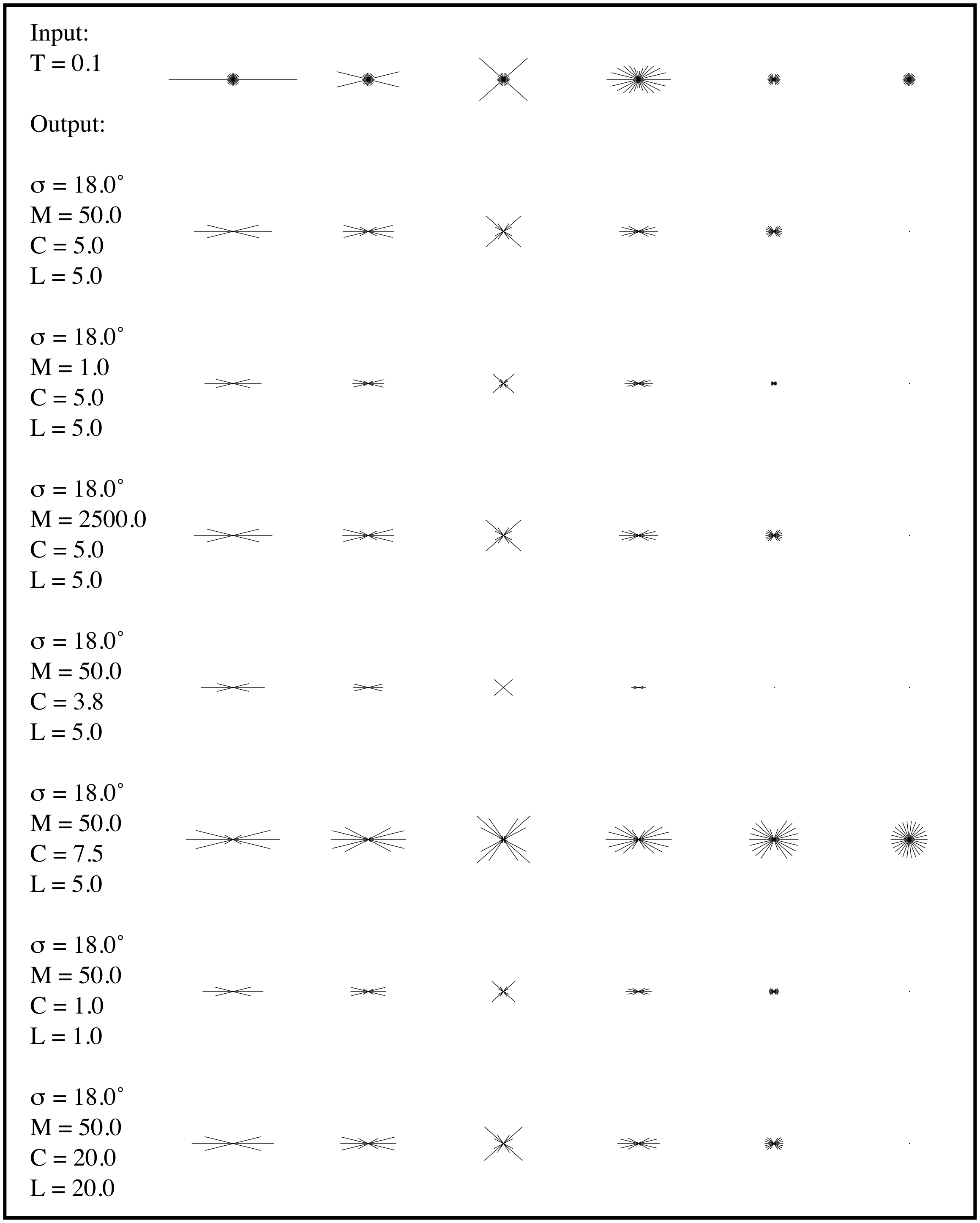


Figure 4-2: Variations of the other parameters. The first row shows the input, same as in Figure 4-1. The second row is the “best” set; the other rows show the effects of increasing or decreasing the dimensionless parameters (M, C, and L).

The tonic activity which is part of all the input patterns used is technically a parameter of the first competitive stage. However, it is important only in the processing of the second competitive stage. Figure 4-3 shows a simulation where the tonic activity of the input patterns is varied. Smaller and larger values than those used previously are tested.

Although larger tonic activities can be helpful for increasing the size of end-cuts, they generally seem to provide too much inhibition. Thus tonic activities larger than 0.1 seem to be detrimental to the output activities of most patterns. Smaller tonic activities do not generate strong enough end-cuts. Therefore a tonic activity of around 0.1 is a good choice.

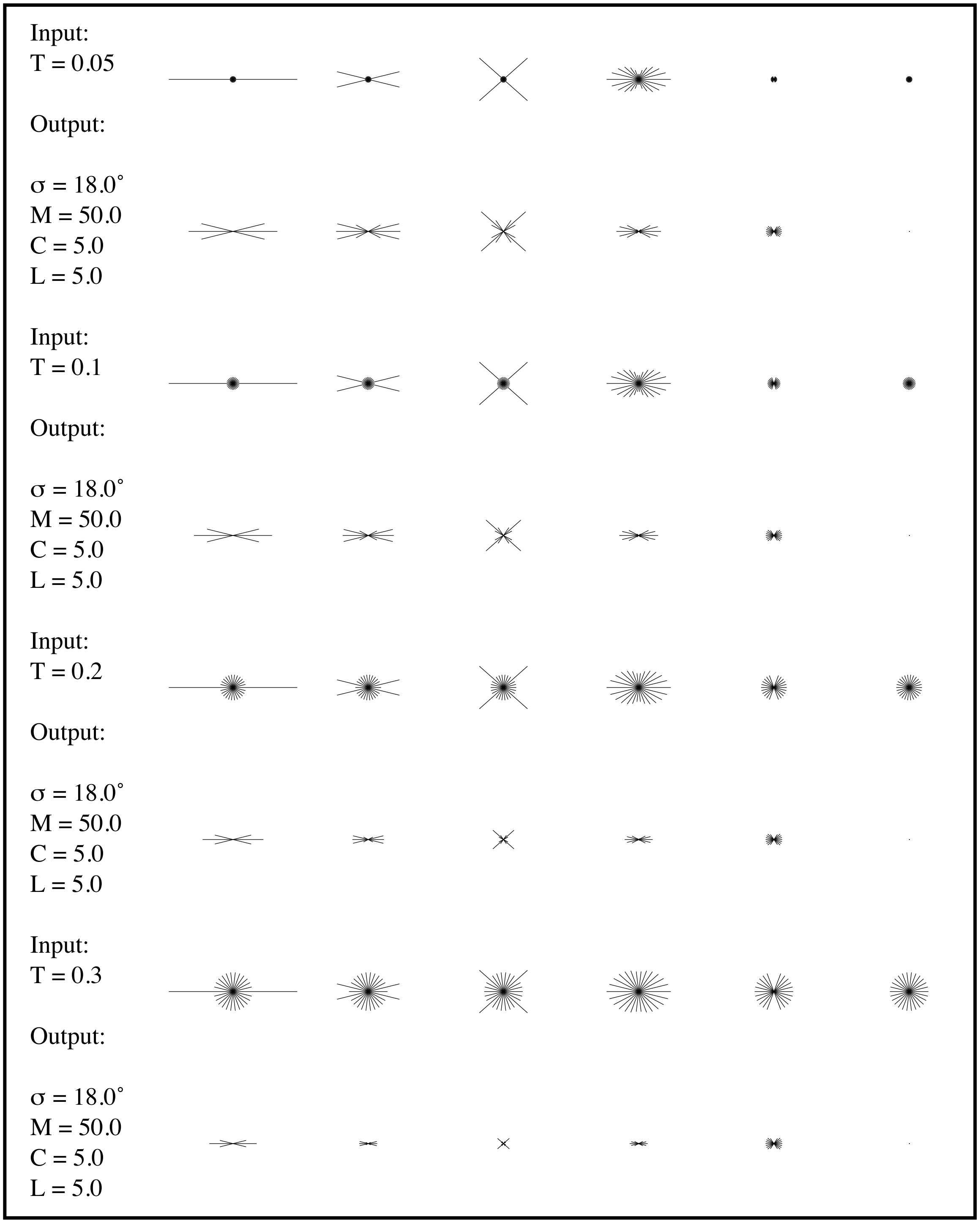


Figure 4-3: Variations in the tonic activity of cells. The inputs have tonic levels (top to bottom): 0.05, 0.1, 0.2, and 0.3. Although larger tonic activities might aid in creating end-cuts, the effect seems largely suppressive.

For all of the parameters, the strongest constraint is provided by the input pattern which is meant to produce an end-cut (pattern #5 in Table 4-1). Even for the best choices for all the parameters, it is still difficult to get a strong end-cut signal out of the second competitive stage.

If necessary, there are changes that can be made which would produce larger end-cut signals. In Figure 4-2, it is clear that when $L < C$ the end–signals are enhanced. The problem, as was discussed above, is that uniform input patterns are not suppressed. However, if we consider what happens to the outputs of this stage, then it becomes clearer why this may be a plausible arrangement. In order to satisfy the postulate on spatial impenetrability (Grossberg & Mingolla, 1985b) the signals from the oriented competition are processed so that activations from the perpendicular orientation are subtracted before entering the long-range completion. This process (see equation B–13) takes care of featural noise suppression, allowing us to have $L < C$.

There is yet another way in which the second competitive stage can be modified in order to produce better end-cuts. Grossberg & Mingolla (1985a, 1985b) originally described the second competitive stage as two sequential operations; the first being the oriented competition and the second being the normalization of activity. As described and simulated above (most similar to the Grossberg & Mingolla (1987) equations), the stage takes as input the activities from the previous stage and in one step performs both competition and normalization.

Therefore, one way to make the second competitive stage more effective would be to "double" the stage by having two of the stages as described previously with the output of one stage being the input to the other. A similar effect can be achieved by making the oriented competition recurrent. Recurrence in this case means that the convolutions of the oriented competition are performed on the *outputs of the oriented competition*. There is no steady–state expression for this type of equation, the solution can only be derived by some type integration of the system of ODE's. It is perhaps the case that *unless* the oriented competition is recurrent, end-cuts can be generated but not normalized. Again, doubling the oriented competition would also allow the normalization can take place.

One of three things can be done in response to these concerns: (1) Nothing. The CC-loop may provide the required "recurrence", and thus normalize end-cuts anyway. (2) Add a second second competitive stage, i.e., doubling it. (3) Make the second competitive stage recurrent, in which case this might as well be the stage that is numerically integrated in simulations (avoiding integrating more than one stage).

Figure 4-4 is an exploration of some of these alternatives. The first row in the figure shows the input; it is the same input which is used in Figures 4-1 and 4-2 as well as in the third row of Figure 4-3. The second row shows the output from a single feedforward stage with the parameters set to previously established ideal values. The third row shows the output after two sequential competitive stages, i.e., "doubling". The fourth through sixth rows shows three to five iterations respectively. Notice how each iteration improves the end-cut. In fact, by five iterations there seems to be a convergence in the form of outputs signifying the presence of a single orientation. The outputs for no orientation, and for two orientations are still distinct.

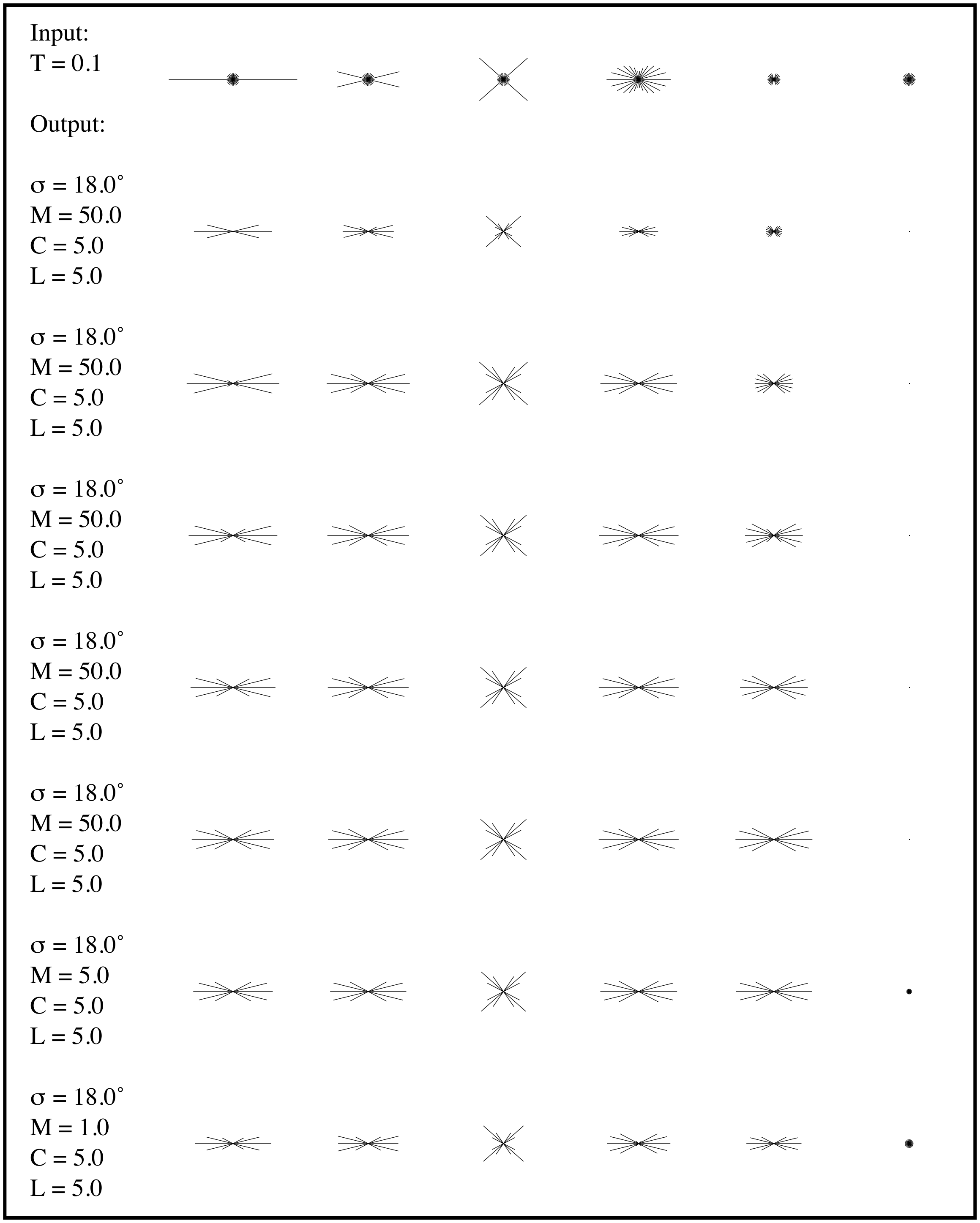


Figure 4-4: Effects of iterating the second competitive stage. Inputs are the same as in Figure 4-1. Row 2 shows previous best parameters, then doubling tripling, 4 iterations, 5 iterations. The last two rows shows the recurrent version.

The last two rows of Figure 4-4 show the outputs of a recurrent second competitive stage, which is described by Equation (4-6) above. Note that in the figure, the M parameter for these two rows applies to the feedback within the oriented competition stage. The M parameter for the input from the previous layer is unity both cases. In the last row, therefore, there is an equal balance between the contributions of the input and the recurrence. However, when the contribution from the input layer becomes more dominant, then the property of featural noise suppression is lost as can be seen in the last row of the figure.

Figures 4-5, 4-6, and 4-7 below show the same parameter variations as Figures 4-1, 4-2, and 4-3, but with the recurrent form of the equation. In general, the recurrent version is not as good at featural noise suppression, but on the other hand it appears somewhat more tolerant of parameters differences. Larger variations of parameters are tolerated, while still yielding the desired results.

In conclusion, it has been possible to establish a narrow range for all of the parameters of the oriented competition, including one parameter of the spatial competition stage (T). It is also worthwhile to note that this specification of all the parameters of the oriented competition was derived without any consideration of what the parameters of the spatial competition might be (except the tonic activity, T). Furthermore, the parameters were derived without the assumption of any particular type of filter to be used as input to the spatial competition; the only assumption is that some type of oriented filter is used. It is now time to analyze the spatial competition in detail, still using a one-dimensional context.

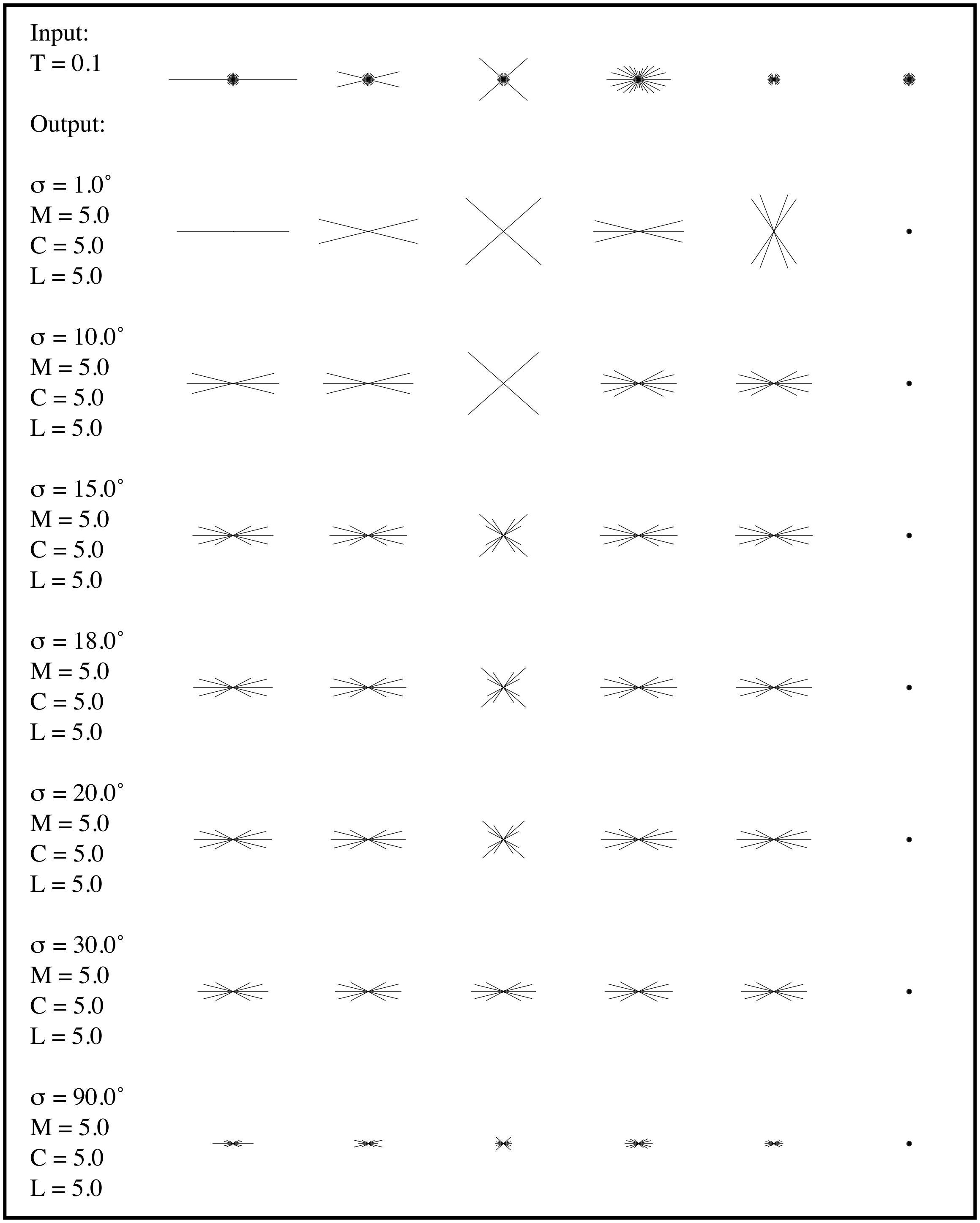


Figure 4-5: A recurrent version of Figure 4-1. The output of the oriented competition given six different input patterns. Seven different widths of the excitatory kernel (σ_c) are used. A slightly wider range, it seems that values in the range 10–20° work best.

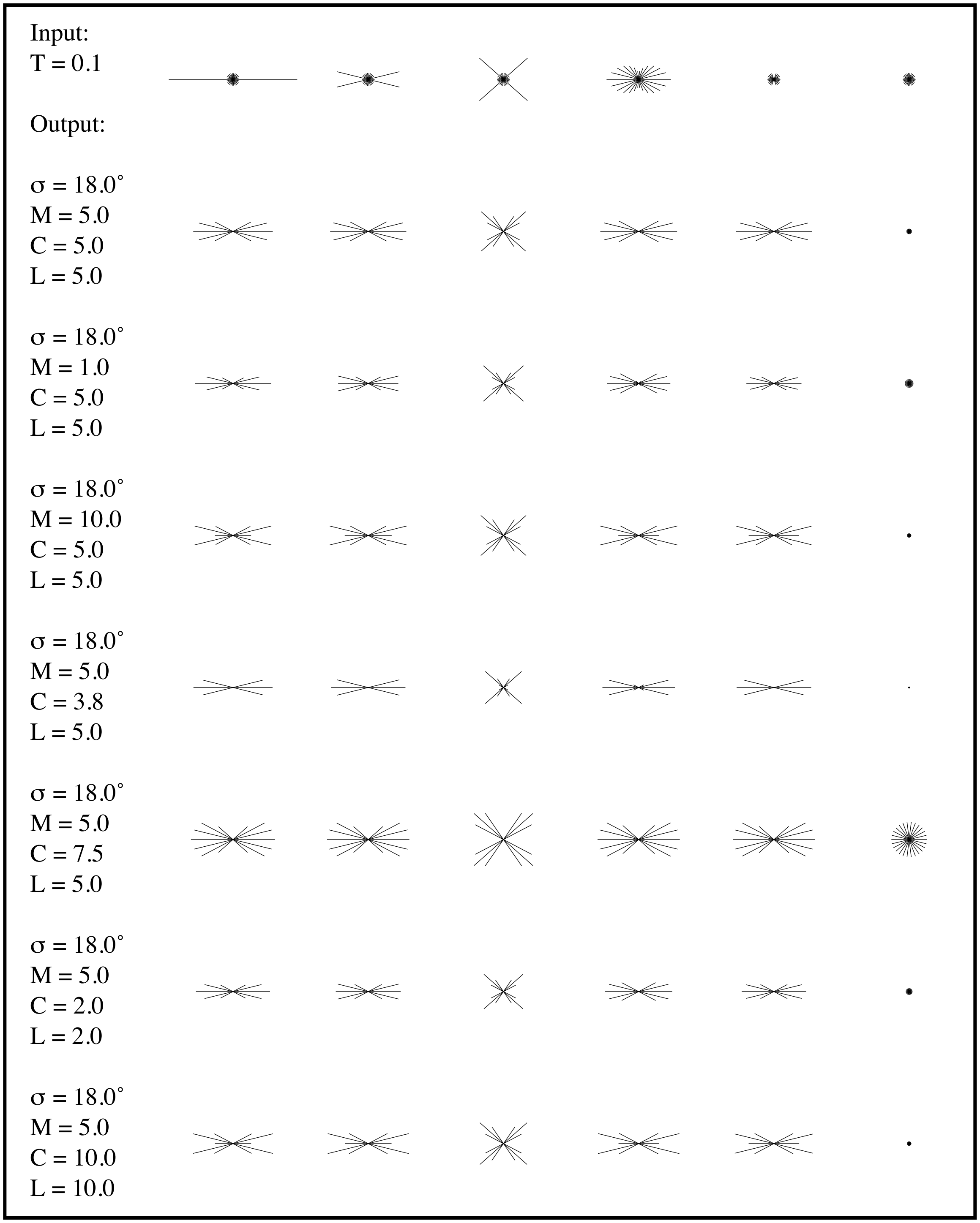


Figure 4-6: A recurrent version of Figure 4-2, variations of the other parameters. The first row shows the input, the second row is the "best" set; the other rows show the effects of increasing or decreasing the dimensionless parameters (M, C, and L), again, this version seems more tolerant of changes in parameters.

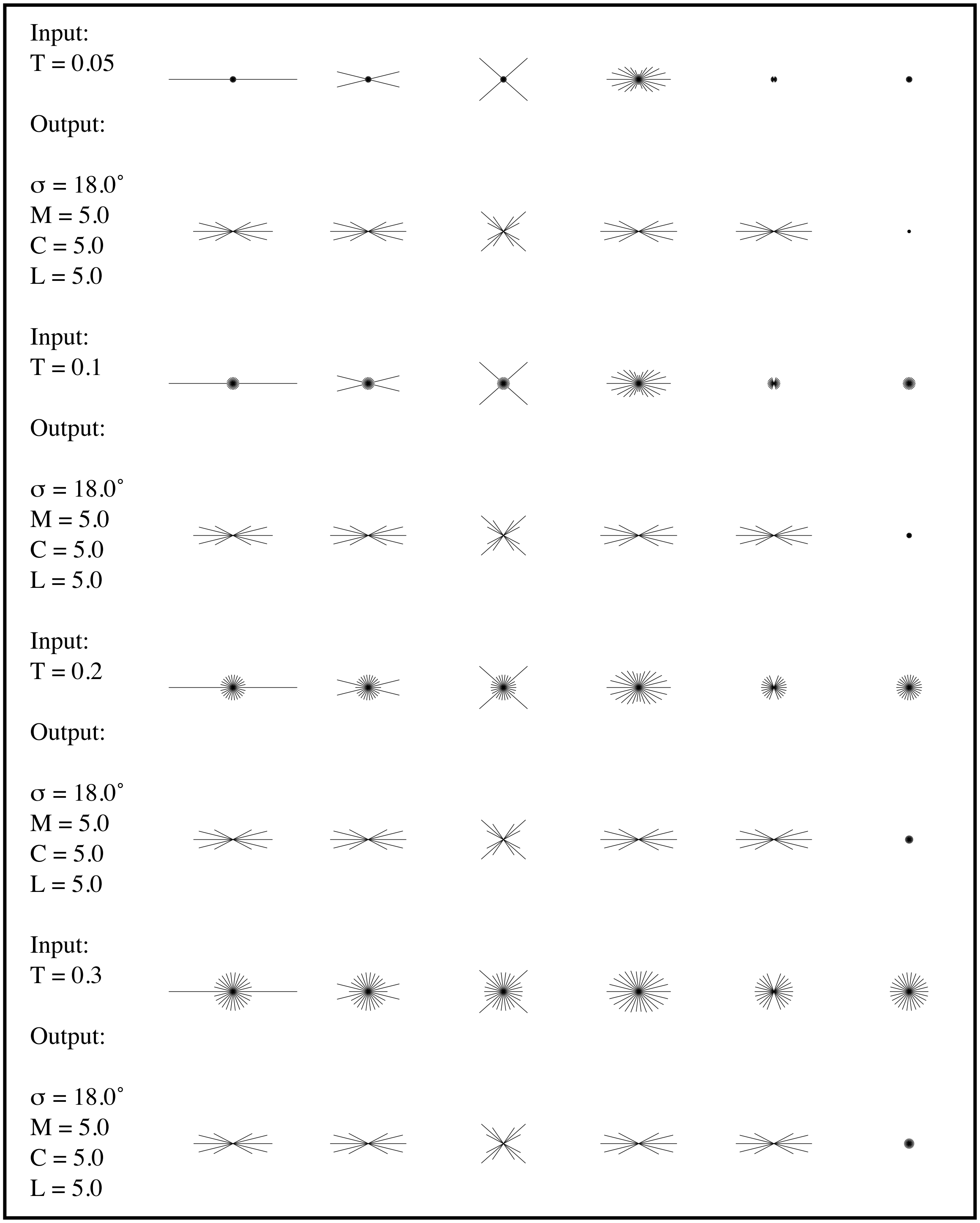


Figure 4-7: A recurrent version of Figure 4-3, variations in the tonic activity of cells. The inputs have tonic levels (top to bottom): 0.05, 0.1, 0.2, and 0.3. The amount of tonic activity appears irrelevant for the recurrent version.

4.4 Analysis of the Spatial Competition

The spatial competition is a stage that takes place *before* the oriented competition (Grossberg & Mingolla, 1985a, 1985b, 1987). The equations describing it are the same as used previously, Equation 4-2. Only the kernels used for the on-center and off-surround competition is spatial rather than oriented.

$$c = N_c \exp(-\left(\frac{\rho}{\sigma_c}\right)^2) \text{ and } s = N_s \exp(-\left(\frac{\rho}{\sigma_s}\right)^2) \;, \tag{4-10}$$

where ρ is the distance of the input to the center of the kernel, and the other parameters are as specified earlier. The parameters σ_c, σ_s, and ρ are specified in units that make sense in the visual field, such as arcmin.

The purpose of the spatial competition is to provide a foundation on which the oriented competition can act, producing end-cuts where they are needed. This is done by inhibition of spatially neighboring nodes with the same orientation. The existence of the spatial competition also seems to explain a number of other phenomena, such as the inhibitory zone of Badcock & Westheimer (1985a; 1985b), but this is secondary to its purpose.

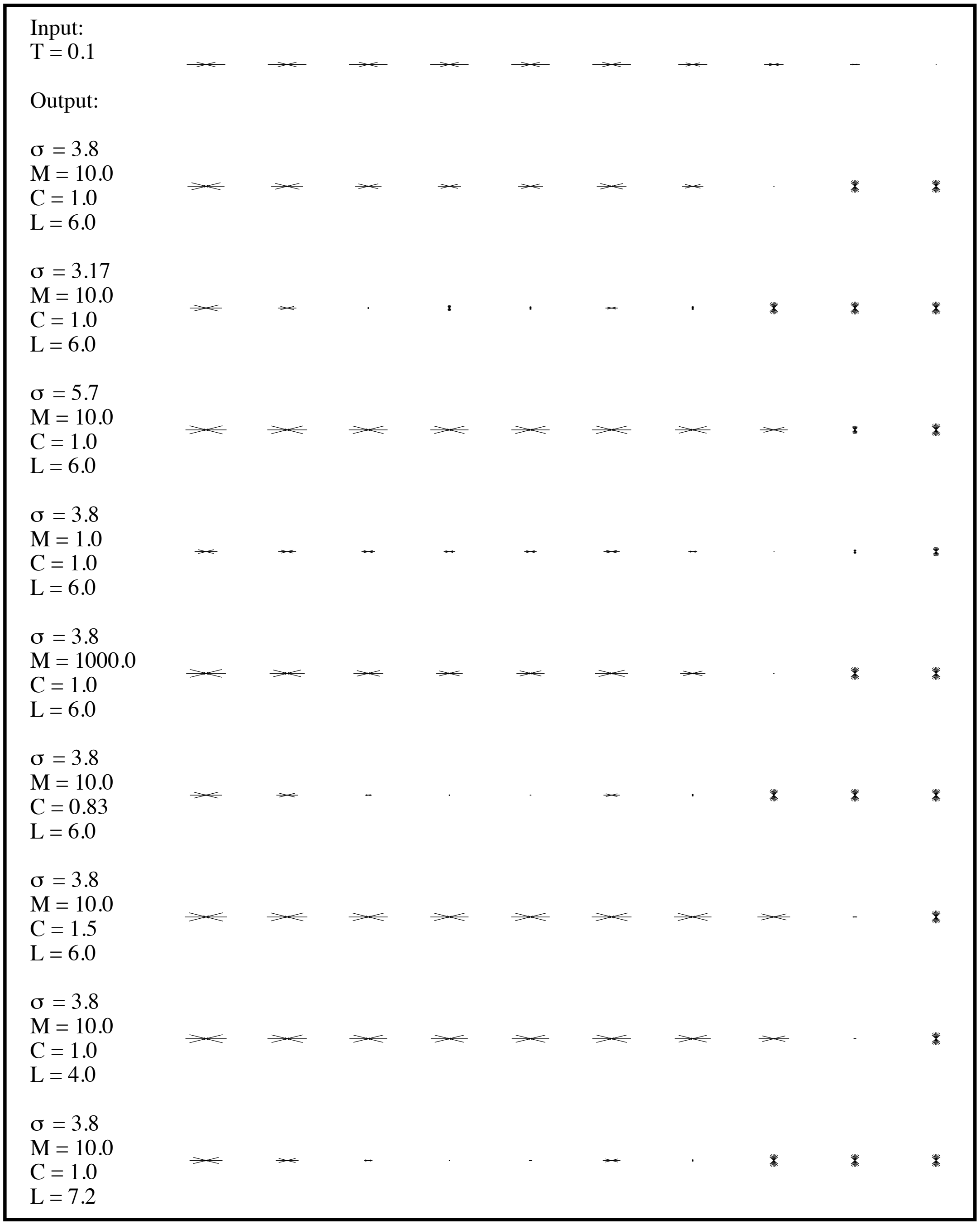


Figure 4-8: Output of oriented competition with input from spatial competition. The first row is the input to spatial competition, which depicts the end of a line, with filter activities slowly tapering off. The second row shows the "best" parameters of the spatial competition; subsequent pairs of rows show the effects of systematically decreasing and increasing the parameters of spatial competition by a factor of 1.2 or 1.5.

The objective in searching for parameters of this stage is to find the proper level of inhibition of neighboring nodes. A balance has to be sought where there is enough inhibition to produce end-cuts, but not so much inhibition as to quench all useful activity. At the end of a stimulus line, filters start becoming less active, this is the site where inhibition should occur in order to produce end-cuts. In the middle of a line, where filters all have almost the same activity, there is inhibition, but not below the tonic activity. These qualitative relationships should also hold true for lines of different contrast. Figure 4-8 shows the output of the output of the oriented competition with the spatial competition. The input is the end of a line with gradually decreasing filter activity.

The parameters of the spatial competition which are left after the dimensional analysis and which can be adjusted according to the criteria above are: M, the scaling of input from "complex cells"; C, the ratio of center to surround; L, the ratio of lower bound to upper bound; and σ_s, the size of the surround kernel. One parameter, T, has already been determined in the simulations of the oriented competition. The size of the excitatory center kernel, σ_c, is assumed to be small.

In the analysis of the oriented competition it was discovered that the ratio of parameters L and C was important for behaviour of the system. It is possible that this ratio may be important for the spatial competition as well. Under close examination, an interesting pattern in Figure 4-8 becomes apparent. In rows 4, 8, and 9, parameters σ_s, C, and L are multiplied by a factor of 1.5, 1.5, and 1/1.5 respectively. Remarkably, the effect on the output in these three rows is nearly indistinguishable. Furthermore, in rows 3, 7, and 10, the same parameters (σ_s, C, and L), are divided by a factor of 1.2, 1.2, and 1/1.2 respectively. Again, the

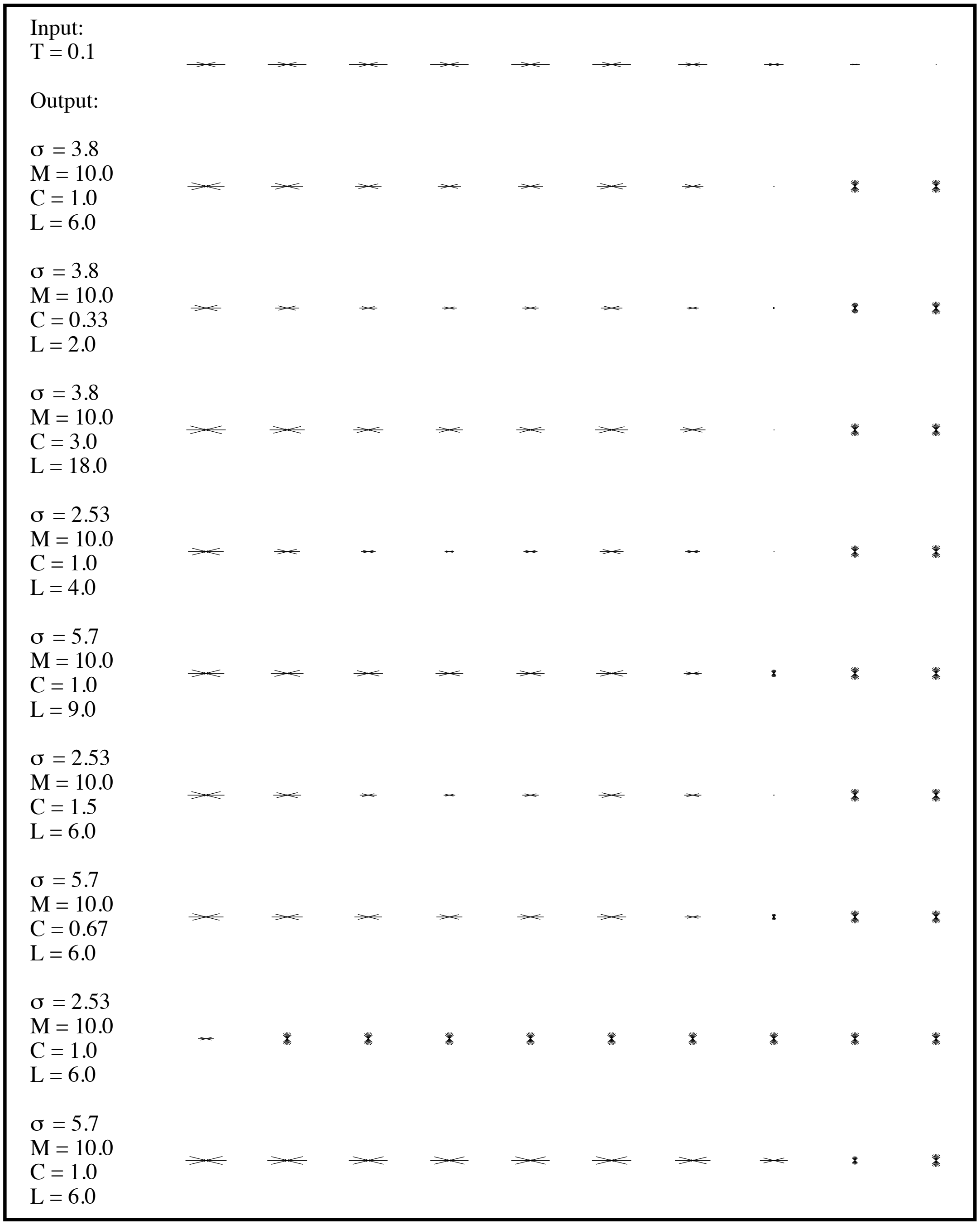


Figure 4-9: Keeping the ratio $L/C\sigma_s$ constant preserves the location of greatest inhibition. The first row shows the input to spatial competition as in Figure 4-8; all other rows show the output of the oriented competition. The second row shows the "best" parameters. The next three pairs of rows keep L/C, L/σ_s, and $1/C\sigma_s$ constant respectively. The last pair changes σ_s only.

output of these rows are very similar. It thus appears that not just a ratio of L to C is relevant, but that the invariant expression is actually $L/C\sigma_s$. This expression changes in exactly the same way when a constant factor is used to modify one of the parameters as done above. The idea that this is a functionally relevant constant is tested in Figure 4-9. The ratios L/C, L/σ_s, $1/C\sigma_s$ are tested one at a time. The last two rows shows the effect of changing only σ_s. Notice that when the $L/C\sigma_s$ ratio is kept constant, the location of the greatest inhibition stays in the same place. When the ratio is not kept constant, and for example only σ_s is changed, this location moves. It moves away from the line as σ_s is larger, and closer to the the line as σ_s is smaller.

A summary of the best parameters based on the criteria discussed in this chapter is presented in Table 4-2. Note that these values are often approximate, and in most cases represent reasonable starting points. The values in the table are the best values for the non-recurrent form of the equation. If a recurrent form is chosen, then these values would probably still work, as simulations indicate that the recurrent form works with a greater range of the parameters.

Table 4-2: Summary of parameter estimates.

Symbol	Description	Value	Model Stage
M	Scaling factor	~10	input to spatial
C	Ratio of on-center to off-surround	~1	spatial
L	Ratio of lower bound to upper bound	~6	spatial
σ_c	Excitatory on-center kernel	small	spatial
σ_s	Inhibitory off-surround kernel	~3.8 “columns”	spatial
T	Tonic activity level	~0.1	spatial
M	Scaling factor	~50	spatial to oriented
C	Ratio of on-center to off-surround	~5	oriented
L	Ratio of lower bound to upper bound	~5	oriented
σ_c	Excitatory on-center kernel	~15-20°	oriented
σ_s	Inhibitory off-surround kernel	large	oriented

While these results are interesting and potentially useful, it is important to remember that they represent simulations performed in isolation with little or no input or feedback from other stages. These simulation also do not use a full two dimensional spatial arrangement of nodes, and no realistic filter model. The next chapter improves on these simulation procedures.

Chapter 5: Hyperacuity Simulations

5.1 Introduction

Hyperacuity, the ability to discriminate stimulus differences smaller than the size of retinal cones, has long held a fascination for scientists interested in vision. Early attempts at understanding this phenomenon focused on the role of the photo-receptors in the eye. As the role of cortical neurons and their responses to various stimuli have become better understood, there has been a shift in trying to understand hyperacuity at the cortical level. Recently, perceptual learning has become an important area of study, and models of hyperacuity incorporating learning have appeared (Sotiropoulos *et al.*, 2011). However, there are still many hyperacuity phenomena which have not been explained.

The models of Wilson (1986) and more recently Sotiropoulos *et al.* (2011) explain how hyperacuity works for simple stimuli. These models are based on simple filters, using several orientations and a number of different spatial scales. When more complex stimuli are used (Figure 5-1)[8], these explanations and models no longer work. For example, when the elements of the stimulus have different contrast polarities with respect to the background, the filters that were able to accurately discriminate hyperacute shifts of same polarity stimuli no longer respond differentially. Or consider masking by a sinusoidal grating; a

8. A demonstration program, including source code, for Mac OS X is available at http://vislab.bu.edu/wp-content/uploads/2011/12/LeviWaugh96.zip

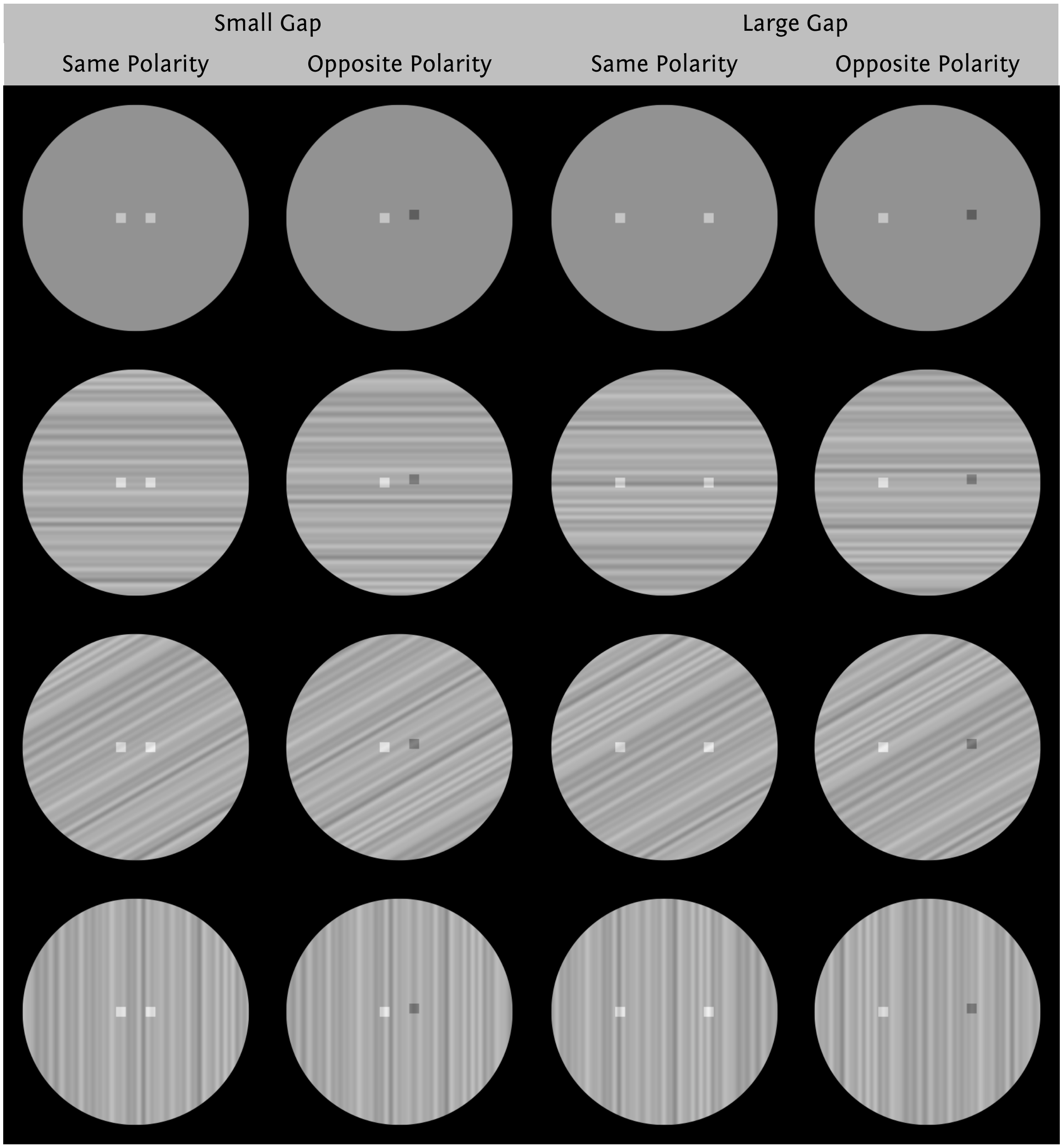


Figure 5-1: Simple and complex hyperacuity stimuli using standard two dot alignment task. Alternating columns show stimuli with the same and with opposite contrast polarity with respect to the background. Alternating columns also show the "no shift" versus "shift". Rows show conditions with uniform background and also with a sinusoidal grating background oriented at 0°, 30°, and 90°, respectively. The top row form the stimuli that are used in Figure 5-15. The threshold elevations observed in the masked conditions below the top row are the stimuli used for Figure 5-12. Redrawn from Levi & Waugh (1996).

filter model cannot successfully dissociate the shift of the stimulus elements from the oriented stimulation provided by the background grating. Even as simple a manipulation as merely increasing the separation between the stimuli (gap size), is not fully explainable with a simple filter bank of different scales. The distinct patterns observed for small gaps and large gaps cannot be explained using just filters of many scales.

In order to explain the results obtained with more complex stimuli, we consider a model of vision that goes beyond simple filter banks. The vision model used is a hierarchical neural network model based on the models of Grossberg & Mingolla (1985a, 1985b, 1987). This vision model builds on a bank of simple filters, but incorporates spatial and oriented competition, edge localization that is insensitive to the direction of contrast, and long-range boundary completion. In addition, divisive normalization is used (Grossberg, 1973; Heeger, 1992; Carandini & Heeger, 1994), allowing the network to deal with both low- and high-contrast stimuli. Variations of this model has been used to explain a range of phenomena (see for example Grossberg, 2003; Kelly & Grossberg, 2000). It is possible to consider this model a variation of the more general class of three stage *filter–nonlinearity–filter* models (Sutter *et al.*, 1989; Graham *et al.*, 1992). Thus while many more models of this type might explain the data dealt with here; only one model is fully described and explored.

Besides the vision model used, there is also a need to connect the vision model with a decision model that links the processing of the vision model with behaviorally appropriate responses. For this we follow Dosher and her colleagues (Petrov *et al.*, 2005; Liu *et al.*, 2010) and use a decision framework with modified Hebbian learning. In this approach, a number of outputs of the vision model are selected, weighted, and fed into a decision node. A similar approach was taken by Sotiropoulos *et al.* (2011).

5.2 Methods

The end result of the simulations reported here, just like in the case of psychophysical experiments involving hyperacuity, is to find the threshold for each specific experimental condition. To find the threshold, multiple stimuli are presented to either the observer or the model, which generate a response for each stimulus. The responses can then be used to compute the threshold. A flowchart of this process is shown in Figure 5-2. There are two main components of the model that take the place of the observer, the *vision* model and the *decision* model. Stimuli are presented to the vision model, which provides a set of inputs to the decision model. The decision model then generates a response for each stimulus. The rest of this section describes each of these steps in detail.

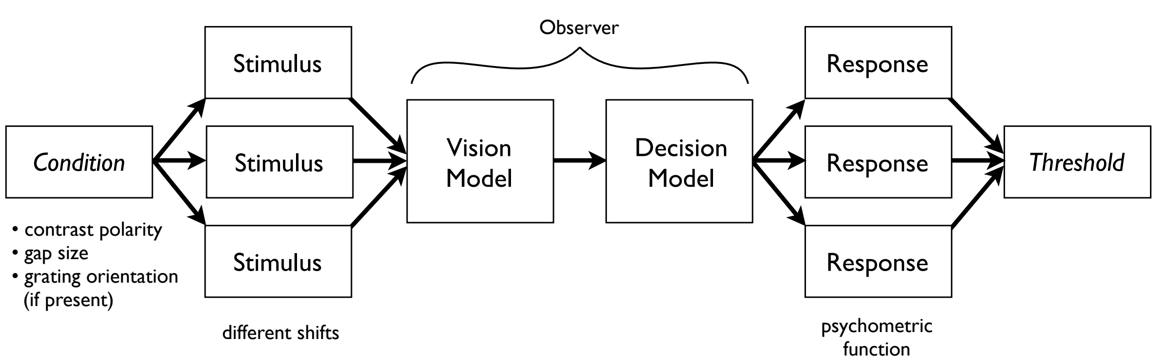


Figure 5-2: Computing the threshold for a condition. This flowchart demonstrates how several stimuli are used for each condition. The stimuli are presented to the vision model (see Figure 5-3), which provides a number of inputs to the decision model (see Figure 5-6), and the decision model generates a response. Any number of standard methods (e.g. Method of Constant Stimuli) can be used to select the stimuli and process the responses into thresholds.

The results in this paper are a set of computer simulations of a model of a part of the human visual system. The model is a variant of the familiar Grossberg & Mingolla series of models (1985a, 1985b, 1987). It is quite likely that other variations that use a filter-rectify-pooling paradigm might also work, although such models might have difficulties with the stimuli requiring long-range completion. Long-range completion is a key feature of Grossberg & Mingolla models.

A key innovation in the current approach is that processing is performed by sampling continuous processing layers. The current work breaks away from traditional image-processing approaches by not using a pixel-grid where computations are expressed by describing the effect of pixels on each other. Instead, the layers of the model are expressed as continuous field equations (see Coombes, 2005; Aubert & Kornprobst, 2006, has some related applications) and computations applied only at a specified set of sampled locations. The set of sampled locations can be a rectangular grid, but it may also be irregular. Because the sampling is implicit, many indices for spatial location and orientation are dropped. By sampling the field sufficiently densely, any side-effects of discrete sampling can be avoided. Note that this discrete sampling is entirely different from *discrete calculus* (Grady & Polimeni, 2010), which is almost the opposite, establishing the operations of calculus on discrete geometrical structures like graphs.

One consequence of writing the equations this form is that the parameters of the model are either dimensionless or have dimensions that make physical sense, such as distance on the visual field [arcmin] or angles and orientations [°]. When parameters are expressed in this manner, it is much easier to compare values for different models or different conditions in a meaningful way. Another consequence of writing the BCS equation in this new way is that they become simpler to understand and one can focus on the essential operations.

5.2.1 Simplifications

In order to simplify the model computations and to make the results more comprehensible, some simplifications have been made. Specifically, (1) only the two highest spatial frequency scales are used, (2) a limited area of the field of vision is simulated, (3) the point-spread function of the eye is incorporated into the filters, (4) binocular interactions are ignored, and, (5) the layers of the model are sampled using a regular grid.

Two spatial scales are used in the simulations. While the evidence from both psychophysical and physiological studies shows that many spatial scales are represented in early vision, the interaction between these scales adds complexity to models of vision, and is not necessary for the types of stimuli considered here. In particular, Waugh & Levi (1995) have used masking experiments to show that spatial frequencies of 9–10 cycles per degree have the most effect on the alignment thresholds of Vernier dot stimuli for separations up to 30 minutes of arc. From these results we can infer that only this spatial frequency scale is necessary. One other coarser scale was added because a “backup” mechanism is needed when the finest spatial scale is completely masked.

Similarly, the whole visual field is not simulated. There are two reasons for this: it would be computationally prohibitive in terms of both time and storage to compute the responses of cells over the whole retina, also, many assumptions would have to be made about extra-foveal processing. The area of the field simulated covers approximately the fovea. It would certainly be possible to simulate para-foveal or peripheral stimuli using the structure of the model presented here, but some modifications would have to be made to account for cortical magnification and the generally sparser representation of different orientations (Westheimer, 2005; Harris & Fahle, 1996) in the periphery.

Especially relevant to hyperacuity is the point-spread function of the cornea. While the precise description of the blurring performed by the cornea is still debated (Geisler, 1984; Ijspeert *et al.*, 1993; van den Berg *et al.*, 2009), there is agreement that an infinitely small point will be blurred to approximately the size of a foveal cone receptor. For the purposes of the simulations reported here, the effects of this point-spread function is absorbed in the specification of the initial filters. Furthermore, any processing performed in the retina or other parts of the visual system that takes place before the initial filter layer is also incorporated into the filters. In addition, the fact that human vision is binocular is completely ignored in the simulations reported here, given that all the psychophysical data is obtained under monocular viewing conditions.

The simulations here use a regularly spaced grid for sampling the layers of the model. The same grid is used for all the layers (except for the sampling of the input, which is much more dense to properly represent hyperacute stimuli), and this is merely a convenience. In principle, the sampling does not have to be on a regular grid at all, it could be jittered, and it could even be different for each layer. But, because of the notation used, the equations and parameters used would still be the same. Testing such variations of sampling would require a different simulation infrastructure than is currently available.

5.2.2 Stimuli & Task

The stimuli used for the simulations are based on the stimuli used by Levi & Waugh (1996), many are shown in Figure 5-1. All stimuli contain two square dots separated horizontally. The dots are 3 arcmin on each side. The separation between the dots ranges from 6 arcmin to 60 arcmin. The dots are presented on either a uniform background or on a compound sinusoidal grating with a specified orientation. The dots have either the *same* contrast polarity with the background (both are brighter), or they have *opposite* contrast polarity (one dot darker and one dot brighter).

The task of the observer in the Levi & Waugh (1996) experiments was to match the perceived relative vertical position of the dots. The dot on the right was either *level* with the dot on the left, or in one of two positions *above* the dot on the left. Feedback was provided after all presentations. The data collected were used to compute the 75% thresholds for each condition using a ROCFLEX procedure (Klein & Levi, 1985). Note that the threshold is therefore expressed as the shift necessary to make the shift or no-shift judgement. As usual, a larger value means that the task is more difficult.

Similarly, the task for the computer simulation is to decide whether the dot on the right is *above* or *below* the dot on the left. Instead of running a large number of noisy trials and estimate the threshold from this data, it is possible to compute the model's analog of a psychometric function directly by simulating a series of shifts, and get the slope directly from these values. This computed threshold is indicative of the difficulty of the task for the model.

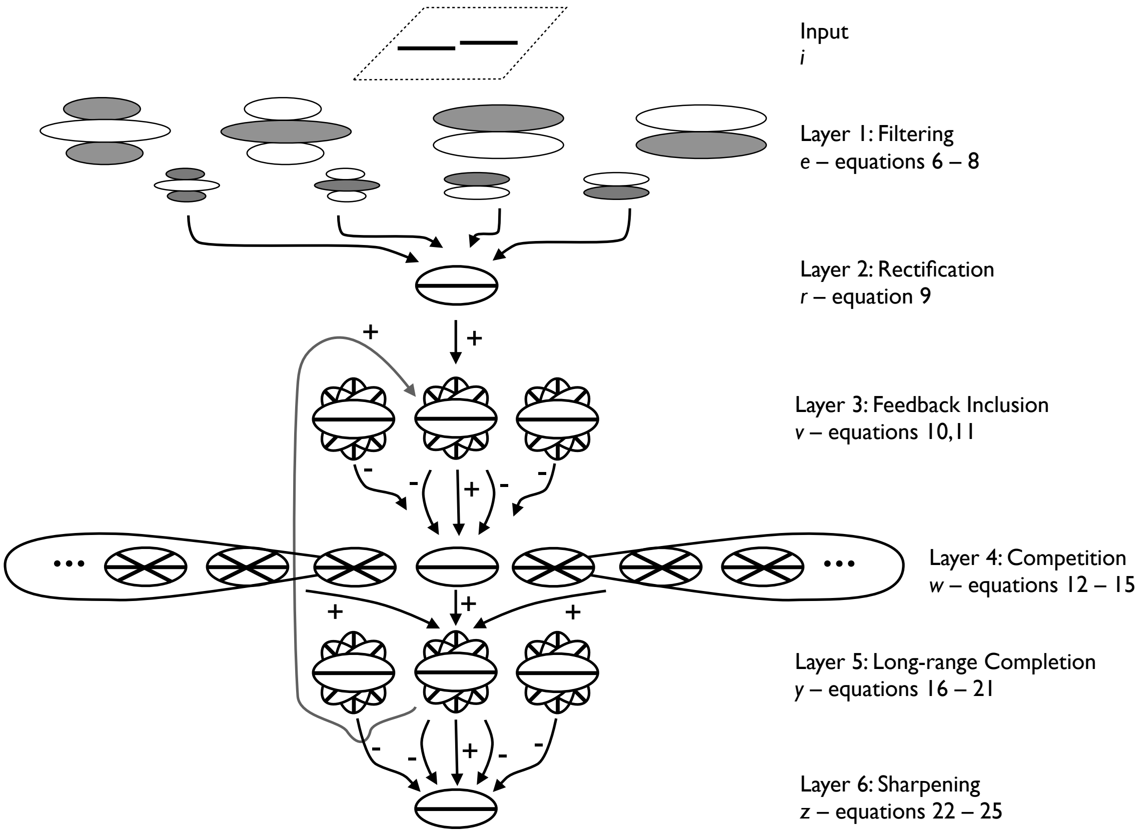


Figure 5-3: The layers of the *vision model* are shown. A layer consists of nodes that sample the activity at a set of locations. Each layer performs a specific type of processing, transforming the activity of the previous layer as indicated by the shapes and the arrows in the diagram. Layer 1 samples the geometrically described input (*i*) using Gabor filters to produce (*e*). Layer 2 combines filters of different phase and different scales into rectified oriented nodes (*r*). Layer 3 combines the oriented filters with feedback to produce (*v*). Layer 4 applies competition, both spatial and oriented, with active nodes inhibiting nearby nodes, to give (*w*). Layer 5 performs long-range completion with bipole kernels that give us (*y*). Finally, layer 6 applies competition again to sharpen the signals and produce the output (*z*). **The figure shows only one orientation at one spatial location; this process takes place at all orientations at all locations.**

5.2.3 Input Model

The stimuli used in the simulations are layered geometrical descriptions, not sampled images representing the stimuli. The descriptions are intensities of types of components (shifted squares, uniform fields, sinusoidal gratings with phase, angle and frequency). This geometrical description is sampled during the filter computation; however, this sampling is sufficiently dense to not deviate significantly from analytically computed valued (which is not possible in the general case).

5.2.4 Vision Model

The model of early vision used is depicted in Figure 5-3. The model consists of six layers of processing. The layers are arranged in a feedforward structure with a feedback connection that goes back to the third layer (v). The processing that transform each layer into the next will be explained in detail using equations.

5.2.4.1 Notation & Conventions

In order to clarify the explanation of the various layers that follow, the notation used in the equations will be explained. When a whole layer is referenced, a bold letter is used (**x**). Scalars and single nodes are denoted by a simple italic variable (x). Note that any indices (p, q, k) of a single node x are always implied and not shown, these indices are integers representing sampling locations; p and q are used for spatial dimensions, and k is used for the orientation dimension (spanning 180°, then wrapping around).

There are some common functions used, such as $\exp(x) = e^x$, also used are

$$[x]^+ = \begin{cases} x & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases} \text{ and } [x]^- = \begin{cases} x & \text{if } x < 0 \\ 0 & \text{otherwise} \end{cases}, \tag{5-1}$$

and finally,

$$\operatorname{sgn}(x) = \begin{cases} +1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0 \end{cases}. \tag{5-2}$$

When a "prime" is used on a node (x'), it means that some type of post-processing has been applied, usually rectification, thresholding and/or scaling; a commonly used function that incorporates all three is

$$x' = W[x - \tau]^{+}. \tag{5-3}$$

Several layers use convolution kernels for processing. These convolutions are written $\mathbf{y} * \mathbf{s}$, which represents a convolution of the layer $\mathbf{y}$ with the kernel $\mathbf{s}$ centered on the location of the current node of the layer. The result is a scalar, and the convolution is performed for every node.

Most of the convolution kernels are a product of three simple Gaussians kernels (two spatial dimensions, one orientation dimension). These kernels are defined by the relation

$$\phi(\sigma_p, \sigma_q, \sigma_k) = \exp(-\left(\frac{\rho_p p'}{\sigma_p}\right)^2 - \left(\frac{\rho_q q'}{\sigma_q}\right)^2 - \left(\frac{\rho_k k'}{\sigma_k}\right)^2) \ , \tag{5-4}$$

where $\sigma_p, \sigma_q, \sigma_k$ define the size of the Gaussians in the three dimensions, p', q', k' are the indices and ρ_p, ρ_q, ρ_k are the corresponding sampling densities. The equations that follow assume that $\rho_p = \rho_q$. The primes on the p', q', k' indicate that they have been rotated to the canonical orientation using

$$\begin{pmatrix} p' \\ q' \\ k' \end{pmatrix} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} p - p_0 \\ q - q_0 \\ k - k_0 \end{pmatrix}, \tag{5-5}$$

where θ is the angle the kernel has been rotated and p_0, q_0, k_0 represent the location of the current node of the layer being convolved.

5.2.4.2 Layer 1, Filtering

The filters used in the model have their origins in the filters of the Wilson (1986) model, which are difference of Gaussian with an extra side lobe in one direction and elongated in the other. One of the problems with this filter formulation is that all filters are even-symmetric, i.e. there are no "edge-detectors", only "line-detectors". In order to have a more general model that will reliably find orientations of edges as well as lines, a different type of initial filter is needed.

In the current model, we have chosen to instead use a Gabor filter equivalent (Daugman, 1980, 1985; Marčelja, 1980; Sotiropoulos *et al.*, 2011). The odd- and even-symmetrical filters are defined by

$$a_{\text{odd}} = \sin(2\pi\omega h)\exp(-h^2/\sigma^2)\exp(-v^2/\alpha^2\sigma^2) \tag{5-6}$$

and

$$a_{\text{even}} = \cos(2\pi\omega h)\exp(-h^2/\sigma^2)\exp(-v^2/\alpha^2\sigma^2)\,, \tag{5-7}$$

where variables h and v are the canonical spatial dimensions of the filter, ω is the spatial frequency (in the h dimension), σ is the width of the filter, and α is the elongation (in the v dimension). Two spatial scales are used. The smaller spatial scale has parameters ω, σ, and α that are tuned to match the highest spatial frequency scale filters of the Wilson (1986) model[9]. The second spatial scale has filters that are double the size, thus ω is half and σ is double. The node activities are defined by convolutions with the input, i.e.

$$e_{\text{even}} = \mathbf{i} * \mathbf{a}_{\text{even}} \text{ and } e_{\text{odd}} = \mathbf{i} * \mathbf{a}_{\text{odd}}\ , \tag{5-8}$$

where $\mathbf{i}$ is the stimulus input and $\mathbf{a}_{\text{even}}$ and $\mathbf{a}_{\text{odd}}$ are the filters defined by Equations 5-6 and 5-7. The relative size of the filters compared to the sampling density is shown in Figure 5-4.

9. The even-symmetric filters match the filters used in Wilson (1986) very closely. These filters have a non-zero response to a uniform input. Other researchers (Spratling, 2010) have used a Gabor filter with an offset such that it always has a zero mean (Lee, 1996).

5.2.4.3 Layer 2, Rectification

The outputs of these nodes are combined such that the orientation signal is maximized (even though the phase information is lost in the process). This is done using a simple magnitude calculation, where the components of the input that were "taken apart" by the sine and cosine components of the filtering layer are now "put together" again. The node values (r) are thus calculated using,

$$r = \sqrt{e^2_{S_1,\mathrm{odd}} + e^2_{S_1,\mathrm{even}}} + \sqrt{e^2_{S_2,\mathrm{odd}} + e^2_{S_2,\mathrm{even}}}\ , \tag{5-9}$$

where e_{odd} and e_{even} are the filter values computed above, S_1 is the smaller scale and S_2 indicates the larger scale.

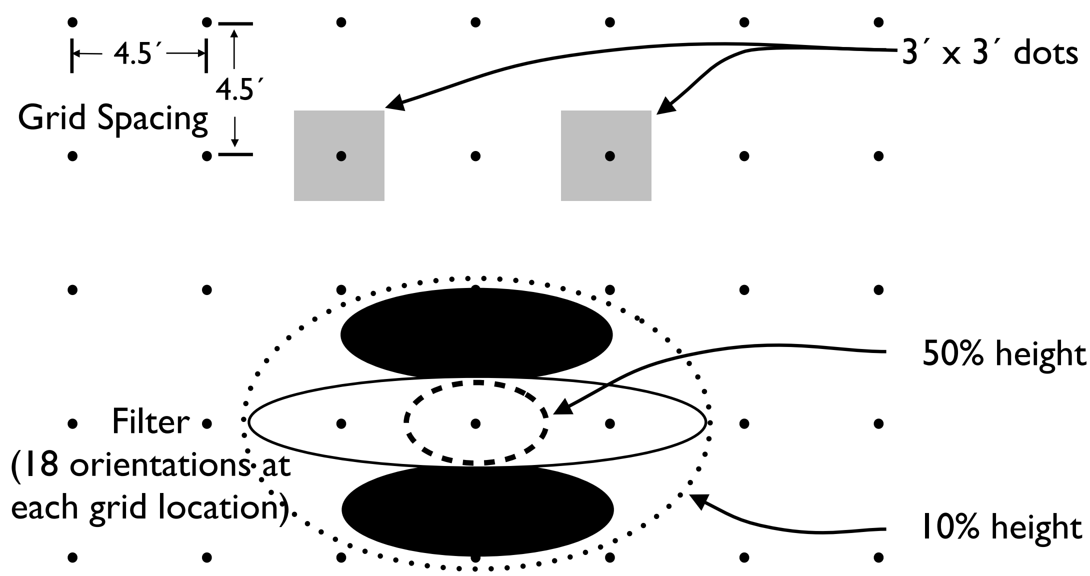


Figure 5-4: A depiction of the sampling grid together with a sample stimuli and a filter at one of the grid locations. A horizontal filter of the smaller scale is shown with the Gaussian half-height curve (50%) as a dashed line; and the 10% curve as a dotted line. The larger scale filters are simply double the size. A sample stimulus with a pair of square dots is shown, accurately reflecting the relative grid spacing and the filter size. The reader should view this at a distance of 11.6 meters for the grid spacing (4.5') to be accurate.

5.2.4.4 Layer 3, Feedback

As shown in Figure 5-3 the nodes in the v-layer combine the bottom-up inputs from the rectified orientation nodes (r) with feedback from the output nodes (y). The activation of v nodes is computed by

$$\frac{\mathrm{d}v}{\mathrm{d}t} = -v + D_v(1-v)\left(\beta r + (1-\beta)y'\right), \tag{5-10}$$

where r is output of the previous layer, y' is the feedback from the long-range completion, D_v is a constant and β is a mixing parameter ($\beta \in [0,1]$). If β is 1.0, then the model becomes purely feedforward with no feedback at all.

5.2.4.5 Layer 4, Spatial and Oriented Competition

The next layer (w) imposes spatial and oriented competition. Nodes representing similar orientations and nearby locations in v will excite nodes in the next layer w. However nodes in v will *inhibit* nodes in the w layer that do not represent similar orientations or that are too far away. This is a common on-center-off-surround kernel, but in three dimensions. The center and surround kernels are defined using Equation 5-4. The center kernel ($\mathbf{c}_w$) is defined by

$$c_w = \phi(\sigma_p^{c_w}, \sigma_q^{c_w}, \sigma_k^{c_w})N^{c_w}, \tag{5-11}$$

where $\sigma_p^{c_w}, \sigma_q^{c_w}, \sigma_k^{c_w}$ define the width of the Gaussians in the three dimensions, and N^{c_w} is the scaling factor that gives the kernel unity volume when summed over all samples. The surround kernel ($\mathbf{s}_w$) is similarly defined by

$$s_w = \phi(\sigma_p^{s_w}, \sigma_q^{s_w}, \sigma_k^{s_w})N^{s_w}, \tag{5-12}$$

again with $\sigma_p^{s_w}, \sigma_q^{s_w}, \sigma_k^{s_w}$ defining the extent of the Gaussians, and N^{s_w} being a scaling factor[10]. The computation of w uses these kernels and is described by

10. While the actual scaling factor is computed by summing the sampled values of the kernel, it

$$\frac{\mathrm{d}w}{\mathrm{d}t} = -w + (1-w)\left[D_w \mathbf{v}' * \mathbf{c}_w + T\right] - (w + L_w)\mathbf{v}' * \mathbf{s}_w \ , \tag{5-13}$$

where v' is the output of the previous layer, D_w and L_w are constants, and T is also a constant representing the minimum tonic activity of nodes. When this is solved for steady-state it yields,

$$w = \frac{D_w \mathbf{v}' * \mathbf{c}_w - L_w \mathbf{v}' * \mathbf{s}_w + T}{1 + D_w \mathbf{v}' * \mathbf{c}_w + \mathbf{v}' * \mathbf{s}_w + T} \ . \tag{5-14}$$

The tonic activity (T) establishes a baseline activity for nodes in this layer. By having a baseline that is higher than the lower bound (L_w), it is possible for activity in the previous layer to inhibit and depress the activity below the baseline in some cases. This is neurally plausible as there are many cells in V1 which have a baseline spiking rate of 5-10 Hz which can be depressed (see for example data in Zhou *et al.*, 2000).

To clarify the meaning of the kernels, when kernels for transforming activity from one layer in to another are discussed, the point of view in this paper is to "look backward", with the reference node being the node in the "to" layer. The kernels describe the weighting of the nodes in the "from" layer, using a position relative to the reference node. The relative position include both spatial position and relative difference in preferred orientations.

---

is possible to estimate the scaling factor for any kernel $\phi(\sigma_p, \sigma_q, \sigma_k)$ using the formula,

$$N(\sigma_p, \sigma_q, \sigma_k) \simeq \frac{\rho_p \rho_q \rho_k}{\sqrt{\pi}^3 \sigma_p \sigma_q \sigma_k \mathrm{erf}(180°/2\sigma_k)} \ ,$$

where ρ_p, ρ_q, ρ_k are sampling densities, $\sigma_p, \sigma_q, \sigma_k$ are the widths of the Gaussians in the three dimensions, and $\mathrm{erf}()$ is the standard error function ($\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$). The error function is necessary, as the Gaussian in orientation space must be truncated as it wraps every 180°, it compensates for the *two* missing tails of this Gaussian.

5.2.4.6 Layer 5, Long-range Completion

The next layer (y) of the vision model performs boundary completion using long-range bipole kernels. These kernels ($\mathbf{b}_z^+$ and $\mathbf{b}_z^-$) have the form

$$b = \mathrm{sgn}(p') \exp(-\frac{p'^2 + q'^2}{\sigma_\rho^2} - \frac{\arctan^2(\frac{q'}{p'})}{\sigma_\theta^2} - \frac{(k' - 2\ \arctan(\frac{q'}{p'}))^2}{\sigma_\varphi^2}) \ , \qquad (5\text{-}15)$$

where p', q', k' are the canonical sampling locations of the "from" node, and $\sigma_\rho, \sigma_\theta, \sigma_\varphi$ are the widths of the Gaussians that make up the kernel. Note that all angular measurements (including k') are expressed in consistent units (degrees have been chosen here). Angles and angular differences must be expressed around zero, e.g. [−90°, 90°], in order to avoid discontinuities in the orientation space. Equation 5-15 describes the *bipole* kernel in the canonical horizontal orientation; which unlike *dipoles* have two positive lobes. These kernels are designed to create smooth contours from fragments. The contour formation process is enhanced by the subsequent feedback. Figure 5-5 shows how the parameters of the bipole kernel relate to the nodes involved. The kernels are used in the node activation,

$$\frac{\mathrm{d}y}{\mathrm{d}t} = -y + g(\mathbf{w}' * \mathbf{b}^+) + g(\mathbf{w}' * \mathbf{b}^-) \ , \qquad (5\text{-}16)$$

or in the steady state,

$$y = g(\mathbf{w}' * \mathbf{b}^+) + g(\mathbf{w}' * \mathbf{b}^-) \ , \qquad (5\text{-}17)$$

where $\mathbf{w}'$ is derived from activity in the previous layer, $\mathbf{b}^+$ and $\mathbf{b}^-$ are the two lobes of the bipole kernel defined by $b^+_{pqk} = [b_{pqk}]^+$ and $b^-_{pqk} = -[b_{pqk}]^-$. It is necessary that both lobes of the bipole are activated, thus the function $g()$, is defined by

$$g(s) = \frac{H[s]^+}{K + [s]^+} \ , \qquad (5\text{-}18)$$

where H and K are constants. Furthermore, y'_k is defined by

$$w'_k = [w_k]^+ - [w_{\perp k}]^+ - \tau_{\mathrm{si}} \tag{5-19}$$

such that signals from layer w are inhibited by orthogonal signals and even then must overcome a threshold (τ_{si}) to implement an outcome known as *spatial impenetrability* (Grossberg & Mingolla, 1985b), so that contours are not formed where strong contours already exist in other orientations. Note also that the output of the y layer is transformed according to

$$y' = M_y[y - \tau_y]^+, \tag{5-20}$$

where M_y and τ_y are constants.

5.2.4.7 Layer 6, Sharpening

The final layer (z) performs sharpening of the long-range completed signals. The mechanism is exactly like the oriented and spatial competition in Layer 4. Nodes representing similar orientations and nearby locations in y will excite nodes in the z layer; and nodes in y will *inhibit* nodes in the z layer that do not represent similar orientations or that are too far away. The center and surround kernels are again defined using Equation 5-4 above. The center kernel ($\mathbf{c}_z$) is specified by

$$c_z = \phi(\sigma_p^{c_z}, \sigma_q^{c_z}, \sigma_k^{c_z})N^{c_z}, \tag{5-21}$$

where $\sigma_p^{c_z}, \sigma_q^{c_z}, \sigma_k^{c_z}$ define the width of the Gaussians in the three dimensions, and N^{c_z} is the scaling factor that gives the kernel unity volume when summed over all samples. The surround kernel ($\mathbf{s}_z$) is similarly specified by

$$s_z = \phi(\sigma_p^{s_z}, \sigma_q^{s_z}, \sigma_k^{s_z})N^{s_z}, \tag{5-22}$$

again with $\sigma_p^{s_z}, \sigma_q^{s_z}, \sigma_k^{s_z}$ defining the extent of the Gaussians, and N^{s_z} being a scaling factor. The computation of z uses these kernels and is described by

$$\frac{\mathrm{d}z}{\mathrm{d}t} = -z + (1-z)D_z\mathbf{y}' * \mathbf{c}_z - (z + L_z)\mathbf{y}' * \mathbf{s}_z \ , \tag{5-23}$$

where y' is the output of the previous layer, D_z and L_z are constants. When this is solved for steady-state it yields,

$$z = \frac{D_z\mathbf{y}' * \mathbf{c}_z - L_z\mathbf{y}' * \mathbf{s}_z}{1 + D_z\mathbf{y}' * \mathbf{c}_z + \mathbf{y}' * \mathbf{s}_z} \ . \tag{5-24}$$

5.2.4.8 Simulation Procedure

The layers are evaluated using the steady-state expressions with initial node values being zero everywhere. In the end, only feedforward processing was used. The feedback path described the above was used for exploratory purposes, but using it did not produce a measurable advantage.

The area of the visual field simulated was 135 arcmin wide by 90 arcmin high. Edge nodes are repeated to infinity. This area corresponded to a grid of sampling locations 31 by 21. The area was substantially larger than any of the stimuli used in order to avoid any edge effects.

5.2.5 Decision Model

The part of the model that connects the output of the visual processing layers to an actual response is a single instar node (Carpenter, 1989), i.e., a summation of weighted inputs that includes a bias term, this is shown in Figure 5-6. The specification of the weights is accomplished using a Hebbian learning paradigm similar to the one discussed by Liu *et al.* (2010) which expanded on the work by Petrov *et al.* (2005), which applied the Hebbian paradigm to orientation discrimination.

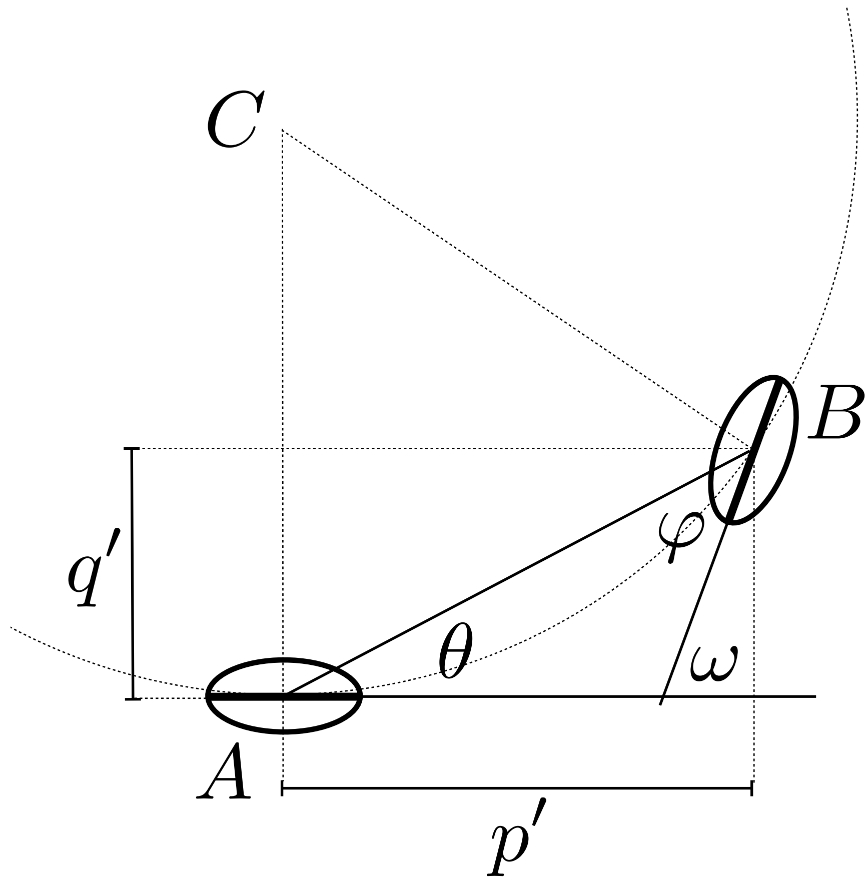


Figure 5-5: A diagram explaining the construction of the long-range bipole kernels. The method is very similar to (Parent & Zucker, 1989; Gove *et al.*, 1995) and see also (Kellman & Shipley, 1991; Field *et al.*, 1993). The center of the bipole is the filter in position A. The strength of the bipole connection from filter B to filter A depends on the relative position and alignment of the filters. Larger coefficients are obtained when the filters are both tangent to a common circle (co-circularity). For the filters A and B this is a circle with a center at C, but note that this circle is different for each filter B. The angle θ is the difference between the ideal (straight line) continuation and the location of the “from” filter B. The angle φ is the angle between the filter B orientation and the line connecting the filters. The kernel coefficient is maximized, and the filter B is tangent, when $\varphi = \theta$. Then the angle ω equals 2θ (since $\omega = \theta + \varphi$). Note that in this figure the alignment is *not* quite optimal ($\theta \neq \varphi$).

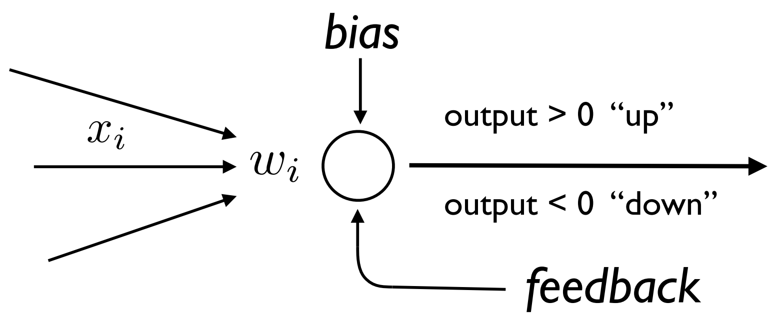


Figure 5-6: The decision model takes input from the vision model and computes a weighted sum that is adjusted by a bias term before a decision is made. In the current context, a positive result indicates a a response of "up", and a negative response indicates a response of "down". The bias term is adjusted by feedback on each response, providing information as to whether responses are correct or not.

5.2.5.1 Selection of Vision Model Outputs

The selection of outputs of the vision model for consideration by the decision model is as important as the structure and function of the decision model. When selecting outputs, choices need to be made in both the *area* of the visual field and which *layer, (or layers),* of the vision model should be used.

In the hyperacuity domain, the focus of activity tends to be a small area, most often in the central visual field. The various conditions reported on in this thesis are no different. Small differences will be registered only in the very central region of the visual field. In terms of the sampling grid specified above, only the center node and the neighboring (8) nodes are relevant. Various combinations of these nodes were tested, with final choice being the central node alone. In terms of which layer or layers of the vision model should be fed into the decision model, again many combinations were tried, but in the end, the output of the long-range grouping stage after another stage of "sharpening" by spatial and oriented competition.

5.2.5.2 Decision Model Equations

The outputs of the vision model, as described and selected above, are now denoted by x_i. The response (u) of the decision unit is computed using,

$$u = \sum_i w_i x_i - b^c, \tag{5-25}$$

where w_i are the weights corresponding to the vision model outputs, x_i are the visual node activities, and b^c is the bias for condition c. This bias can be learned using slow learning, or it can be computed from a few conditions. The response (u) is transformed into the range $(-1, +1)$ using the standard logistic sigmoid function,

$$u' = \frac{2}{1 + e^{-\gamma u}} - 1 \text{ (or equivalently, } u' = \frac{1 - e^{-\gamma u}}{1 + e^{-\gamma u}}\text{)}, \tag{5-26}$$

where γ is an arbitrary scaling factor.

5.2.5.3 Modified Hebbian Learning

The weights (w_i) are adjusted using the feedback (f) using modified Hebbian learning along the lines of $\Delta w_i = f(x_i - w_i)$. The actual weight adjustment also limits the weights to the range $[-1, +1]$ using

$$\Delta w_i = \epsilon(1 - w_i)[f]^+ + \epsilon(w_i + 1)[f]^-, \tag{5-27}$$

where ϵ is a small learning rate parameter (Liu *et al.*, 2010).

The weights were learned initially, and then kept constant for all conditions. Only the bias term (b^c) was adjusted for each condition. The initial learning for the shift sensitivity is equivalent to learning orientations 30° on either side of the horizontal, or for a gap of 6 arcmin this is a shift of 5 arcmin (the dots are 3 arcmin square).

One way to think about the actual computation in simple terms is that it ends up being close to a simple difference between the final activations at 30° to either side of the horizontal, as in,

$$u = x_0^{+30^\circ} - x_0^{-30^\circ} \tag{5-28}$$

but it is important not to forget about the bias needed for each condition, b^c,

$$u = x_0^{+30^\circ} - x_0^{-30^\circ} - b^c. \tag{5-29}$$

5.2.6 Threshold Computation

The threshold for a specific condition corresponds to the inverse of the slope of the response function (psychometric function analog) as the vertical shift of the right dot is varied. If the slope is shallow, then the threshold is large; and if the slope is steep, then the threshold is small. The threshold is estimated by computing the response at a shift of zero arcmin and at a shift of 0.5 arcmin, the output is computed and the slope estimated; the inverse yields the threshold. This method closely corresponds to the method used by Waugh *et al.* (1993) for computing psychophysical thresholds. The threshold computation can be mathematically described as the slope of u' evaluated at $u = 0$ ($u'(u = 0)$ or $u'|_{u=0}$).

5.2.7 Model Calibration

As described, the model generates a response for every stimulus. Using a method of constant stimuli (or some other method suitable for computing thresholds), thresholds can be computed for each condition, and these thresholds can be compared across conditions. However, because the model computed responses use an arbitrary (but internally consistent) scale, the model-produced threshold values need to be scaled before they can be compared to psychophysically measured thresholds.

In the case of the results reported here, we have chosen *one* experimental condition, and the threshold for that condition is scaled to the actual threshold measured experimentally. The scale is set by adjusting parameter γ in Equation 5-26 above. This same scale is then used for *all* computed thresholds. The chosen experimental condition is the two-dot Vernier alignment task with a gap between the dots of 24 arcmin, and with the dots having the same contrast polarity with the background and no background grating.

5.3 Results

Using the methods described above, simulations of many types of input conditions were performed. Here we describe the results of some of those simulations and how they explain some challenging data from (Levi & Waugh, 1996). Many of the figures used to present the output simulations use a "needle-plot". An example of a needle-plot and a guide to interpretation is shown in Figure 5-7. The figure represents one spatial location, but this spatial location has a number of nodes (18) coding different orientations. The activity of each node coding a different orientation is represented by a line of that orientation, and the length of the line is proportional to the activity of the node.

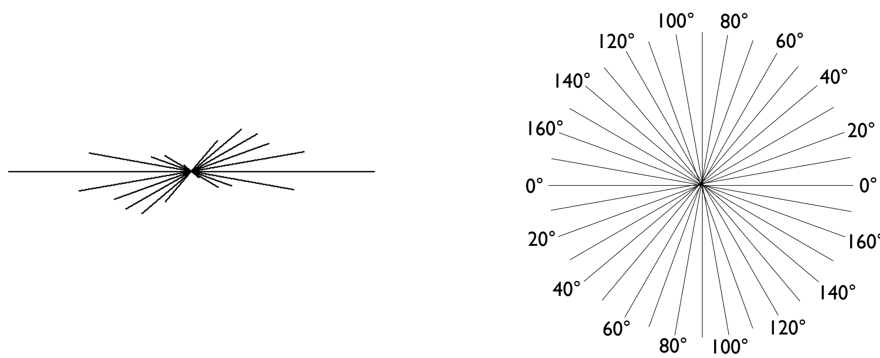


Figure 5-7: (left) A needle plot of one spatial location. For each node coding a distinct orientation, there is a line that represents that node. The orientation of the line matches the orientation coded by the node, and the length of the line is proportional to the activity of the node. (right) All of the represented orientations are shown; they are equally spaced.

5.3.1 Small Gap, Same Polarity

The first case to consider is where the gap is small (6 arcmin) and the dots have the same polarity with respect to the background. Needle plots of the activities of different layers of the model are shown in Figure 5-8 below. As can be seen, the filter activity extends around the two dots. The small filters extend about four sampling intervals (about 18 arcmin), and the large filters extend double that, about eight sampling intervals (about 36 arcmin). The larger filters are much larger than the stimulus in the small-gap condition, and while they cover the dots completely, they are not able to discriminate small shifts very well because of lack of precision.

The plots display the activity of nodes at different spatial locations and orientations. The node activities are grouped by spatial location, and the orientation of the lines correspond to the orientation of the node it represents. The plots in Figure 5-8 show only the central 15 by 15 arcmin area of the visual field. The stimulus dots are represented by yellow squares (for positive contrast with the background) or red squares (for negative contrast with the background). The black lines in the plots show positive activation of the node, red lines show negative activation. Negative activation is generally not passed on to the next layer, but is shown as it helps in understanding the mechanisms of the model. In addition, the negative activation may in some cases correspond to the activations of inhibitory interneurons, and may therefore be helpful in understanding possible physiological correlates of the model.

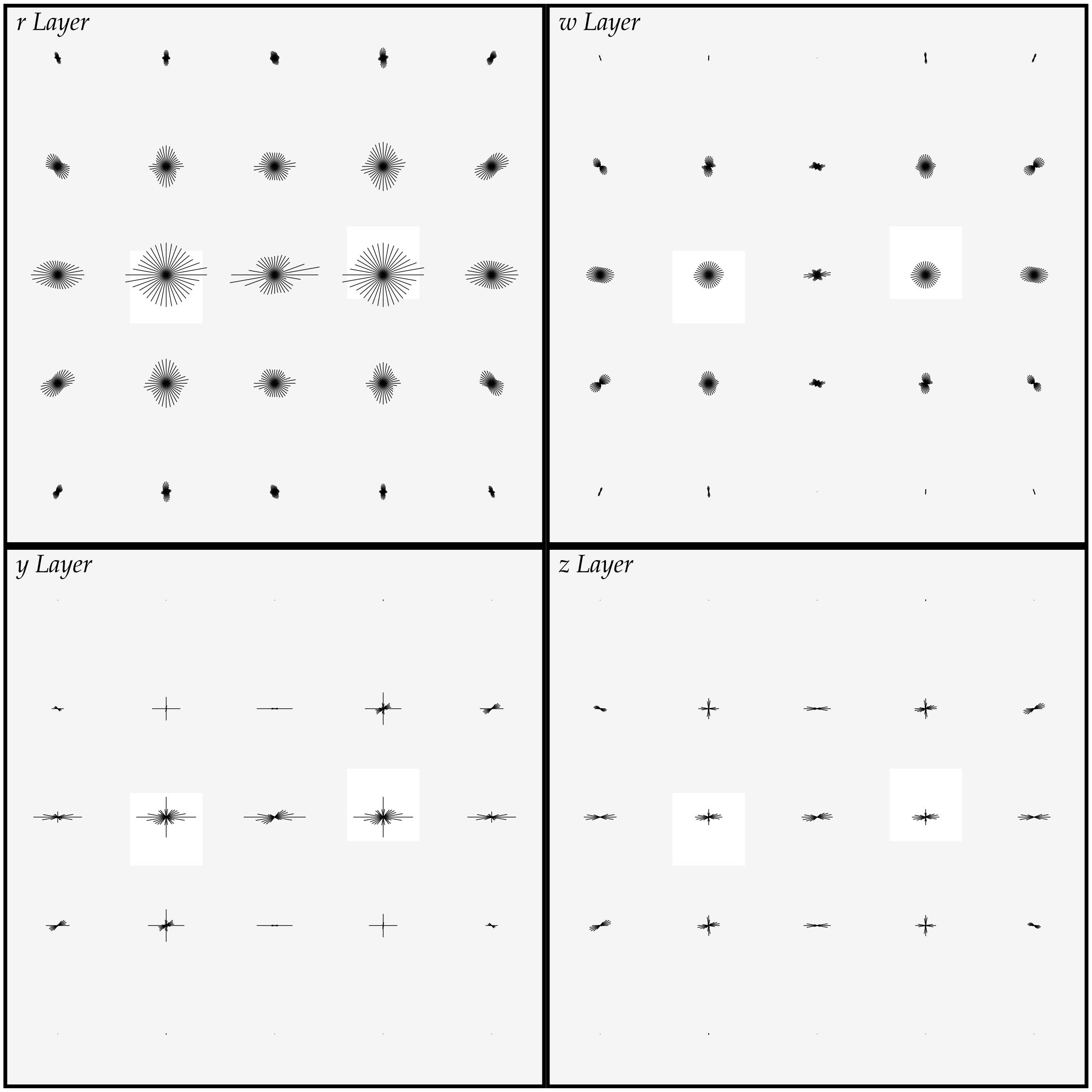


Figure 5-8: Responses of nodes in layers *r*, *w*, *y*, and *z* to same polarity dots separated by only 6 arcmin. The shift is 1.0 arcmin. These plots show only the central 15 by 15 arcmin of the visual field simulated. The shift is clearly visible and represented in all the layers of the model by the near-horizontal orientations of the node in the center.

The stimulus used for this case has the right dot shifted up by 1.0 arcmin. This is evident in the responses of all the layers in the figure, but is perhaps strongest in the response of the filter layer (r), but is also present in the subsequent layers. The key to detecting a vertical shift in the position of the dots is not to examine the signal of the node representing the *horizontal orientation*, but instead to look at the nodes representing 30–40° on *either side* of the horizontal. The *change* in strength of the activation for a small shift is what matters, and thus the activation of the horizontally oriented node is irrelevant.

5.3.2 Small Gap, Opposite Polarity

When the contrast polarity of the dots with respect to the background is not the same, the the responses of the initial filters do not represent the shift of the dots in the same manner. Figure 5-9 shows the responses of the model to this stimulus. The filter responses are split into a bimodal distribution with maximal activities corresponding to where the lobes of the odd-symmetric filters match up with the opposite contrast dots. After spatial and oriented competition, this bimodal distribution although diminished, remains. The long-range linking and subsequent competition together create a unimodal distribution that represents the shift of the dots in the stimulus much in the same way as in the case of the same polarity dots described above.

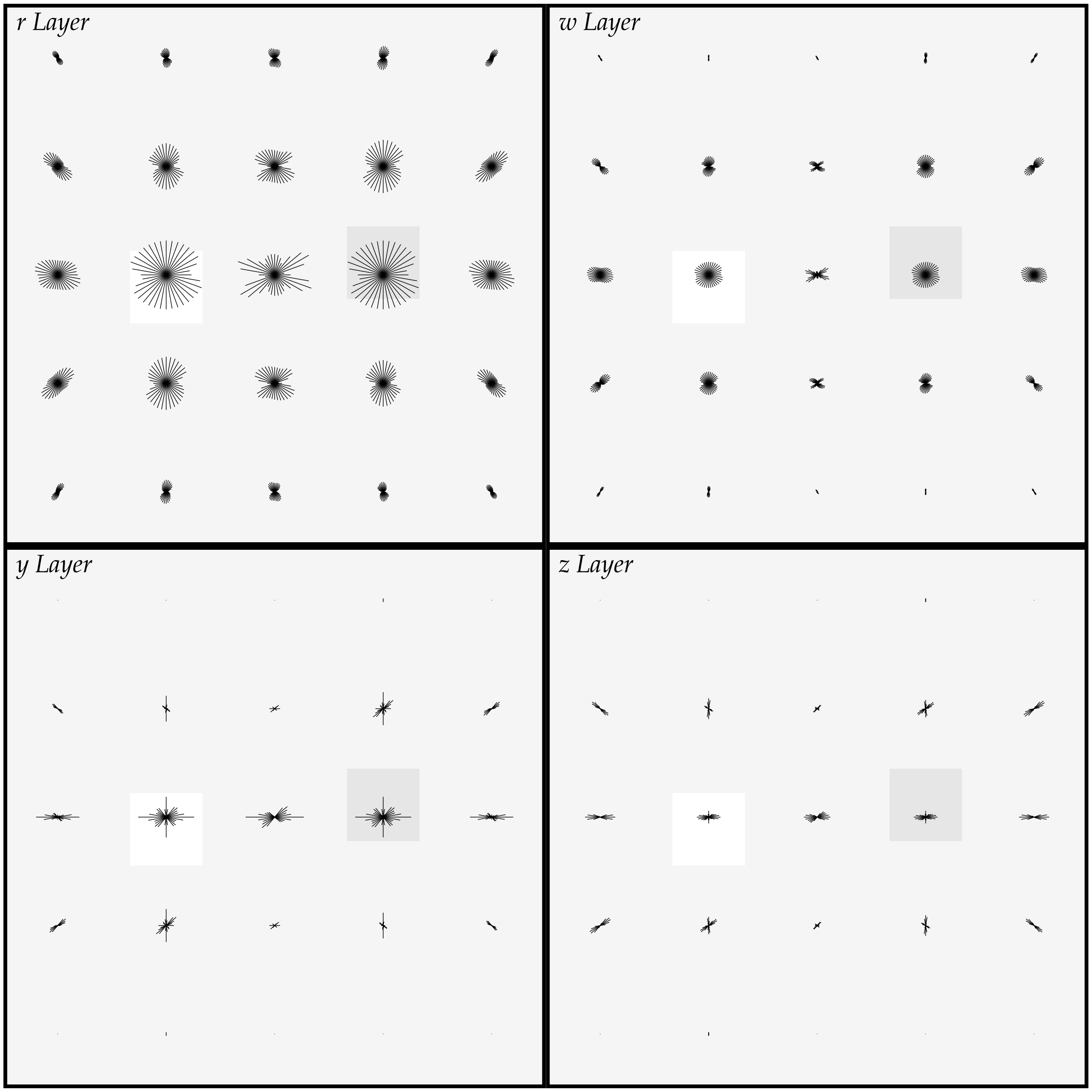


Figure 5-9: Responses to opposite polarity dots separated by 6 arcmin. Notice how filter (r) and competitive (w) layers have a bimodal distribution, but that the long-range linking (y) and sharpening (z) work to create a unimodal distribution that encodes the shift of the dots.

5.3.3 Large Gap, Same and Opposite Polarity

When the gap between the dots is increased, the expectation is that the small dots are still best represented by the smaller spatial scale, but that the long-range linking mechanism is able to link them up and accurately represent any shift in the dots.

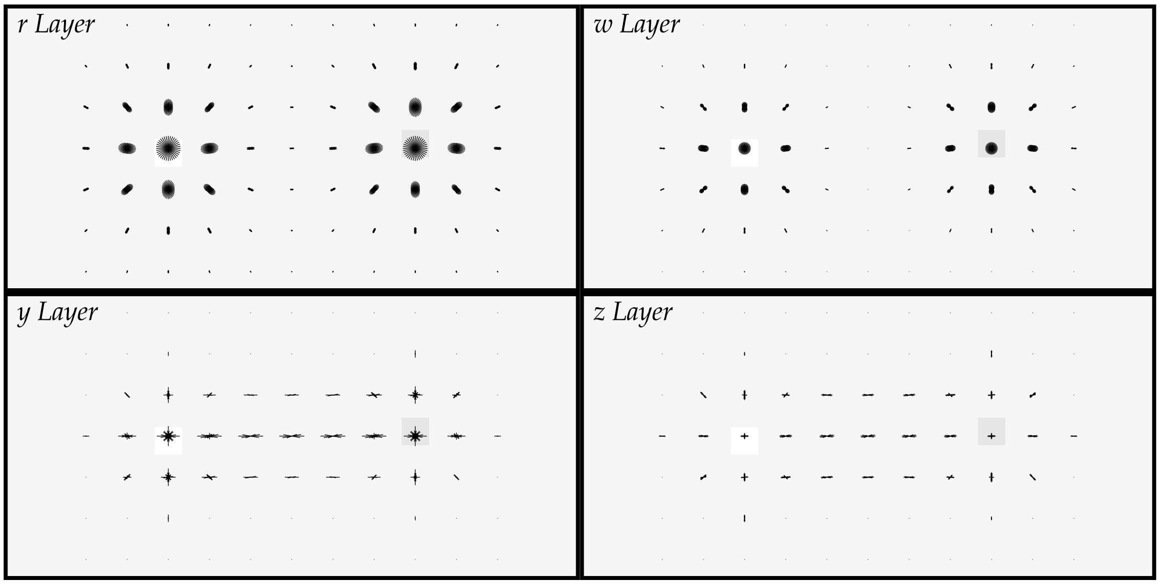


Figure 5-10: Needle plots of the responses of the different layers of the model when presented with two dots with a large gap. In this case the dots have opposite polarity with respect to the background. The two dots do not interact in the filter (r) or competitive (w) layers. The long range completion layer (y) and the additional sharpening layer (z) link the two dots, and is able to detect small shifts orthogonal to the orientation of the gap. The area of the visual field displayed in these plots is 40 arcmin by 20 arcmin.

Figure 5-10 shows the activities in the layers of the model when the gap size is 24 arcmin. As expected, the dots are well represented by the smaller filters, and there is no interaction between the dots. The long-range linking nicely creates a representation of the connection between the dots as can be seen in the bottom two panels.

If the dots have the same polarity, then the plots should be quite similar. Only the filter layer should be affected by the difference in polarity since there is no interaction between the dots. In fact, Figure 5-11 shows the center of the z layer in both cases, the same polarity condition on the left, and the opposite polarity condition on the right. The plots are indeed visually identical, which demonstrates that the model is working as expected.

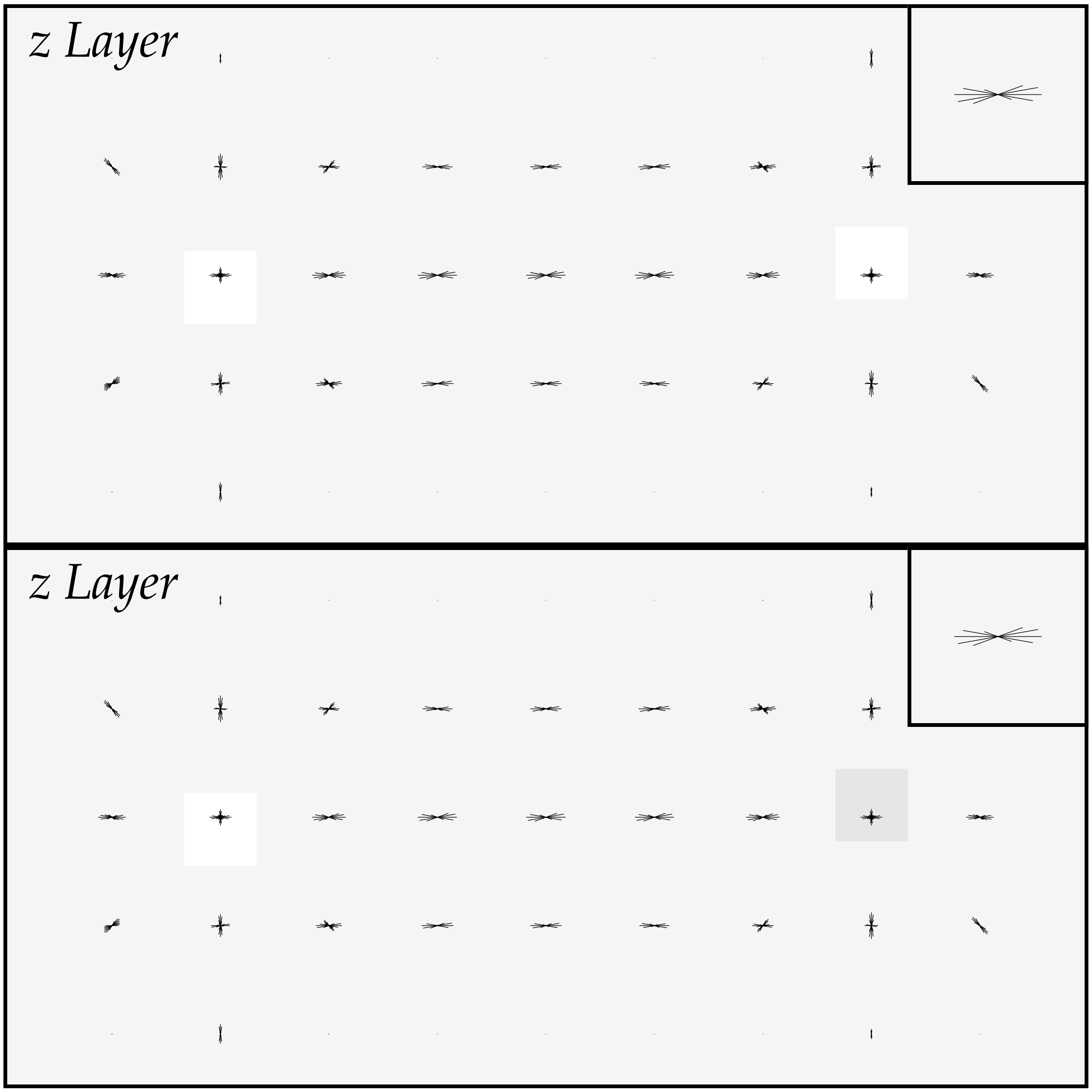


Figure 5-11: The z layer when presented with dots that have the same polarity (top) and opposite polarity (bottom) with respect to the background. Despite being generated by stimuli with different contrast polarities, these fields are indistinguishable. Even the center nodes magnified (see inset in top-right of each plot) look identical. The view shown in these plots is 30 x 15 arcmin.

5.3.4 Levi & Waugh Figure 4: Thresholds vs Gap Size

The main results presented in this paper are shown in Figures 5-12 and 5-15 below. The plots on the left side show data from (Levi & Waugh, 1996) and the right side shows data from the simulations. The first of these is Figure 4 from Levi & Waugh, which shows how thresholds increase with increasing gap size. There are two interesting effects shown in the figure. One effect is in the small-gap regime, and where same polarity stimuli have thresholds significantly lower than opposite polarity stimuli. The other effect is in the large-gap regime, where the curves are almost identical for same and opposite polarity stimuli, and thresholds are increasing steadily. The explanation of the effects in terms of the model can be summarized in the following ways:

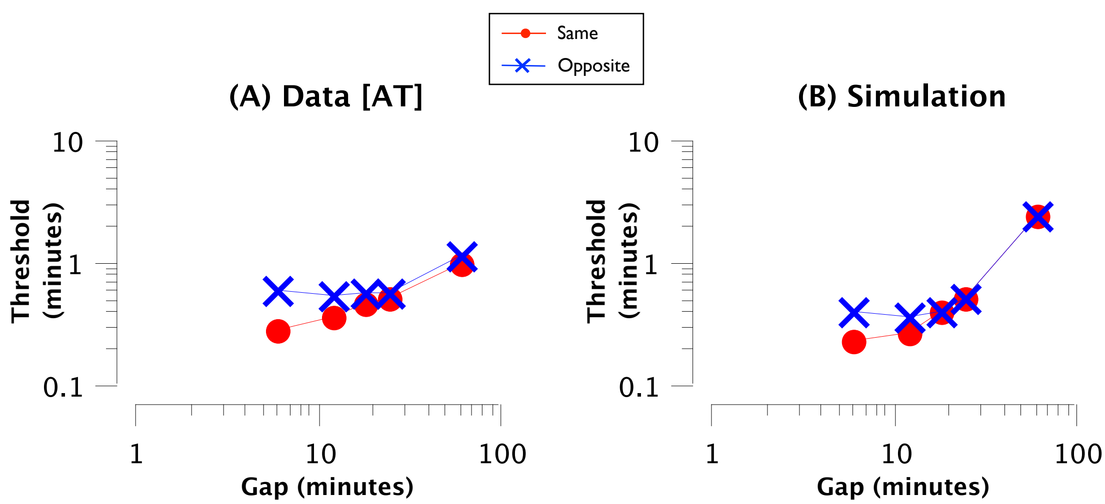


Figure 5-12: Threshold vs Separation of a two dot Vernier stimulus. The blue line (crosses) denote opposite contrast stimuli and the red (circles) denote same contrast stimuli. There are two interesting effects shown. One effect is in the small-gap regime, and where same polarity stimuli have thresholds significantly lower than opposite polarity stimuli. The other effect is in the large-gap regime, where the curves are almost identical for same and opposite polarity stimuli, and thresholds are increasing steadily.

The small-gap domain is dominated by the effects of the filters, which are large enough to cover both dots. The filters that are oriented horizontally or near horizontally respond well to the same-polarity stimuli. Filters at the center of the visual field also respond to the opposite-polarity stimuli, but the distribution of responses is quite different from the same-polarity case. The filters that respond the most to opposite polarity stimuli are not the same as the filters that respond the most to same polarity stimuli. The oriented competition has a comparatively greater effect on the filter responses to opposite polarity stimuli, and thus the thresholds are lower for this condition. The large-gap domain is dominated by the effects of the long-range completion, and this is stimulated equally well by the same and opposite polarity dots. These effects were discussed in detail above.

5.3.5 Sinusoidal Grating Background (+30°)

When the stimulus dots are presented on a sinusoidal grating background rather than on a uniform background, the responses of the model layers will be dominated by the sinusoidal grating. It has been found psychophysically that such gratings interfere with human observer's ability to detect alignment thresholds (Levi & Waugh, 1996). In fact the pattern of interference is especially compelling in supporting the idea that there are other mechanisms than filtering that need to be included in a vision model.

Before looking closely at those patterns, however, consider the case of two same polarity dots, separated by a small gap (6 arcmin), on a sinusoidal grating background. The responses of the layers of the model are shown in Figure 5-13, and the majority of activity in the plots is a result of the oriented sinusoidal grating. The layer after the spatial and oriented competition (w) is most interesting, as it is able to discount much of the activity of the grating and actually represent the pattern of the two dots. This can still be seen after the long-range completion (y), where there is more activity in the horizontal (and vertical!) orientations at the locations of the dots, and also in between the dots.

It is necessary to remember again that the key pattern to observe is not merely the existence of the pattern of dots in the "sea" of the oriented grating, but the *change* in this pattern as the shift changes. To this end, Figure 5-14 shows the z layer when the shift is zero (shift = 0) for the same condition. On the right, the figure also shows the actual difference in node activations between the z layer with *shift*=+1 and *shift*=0.

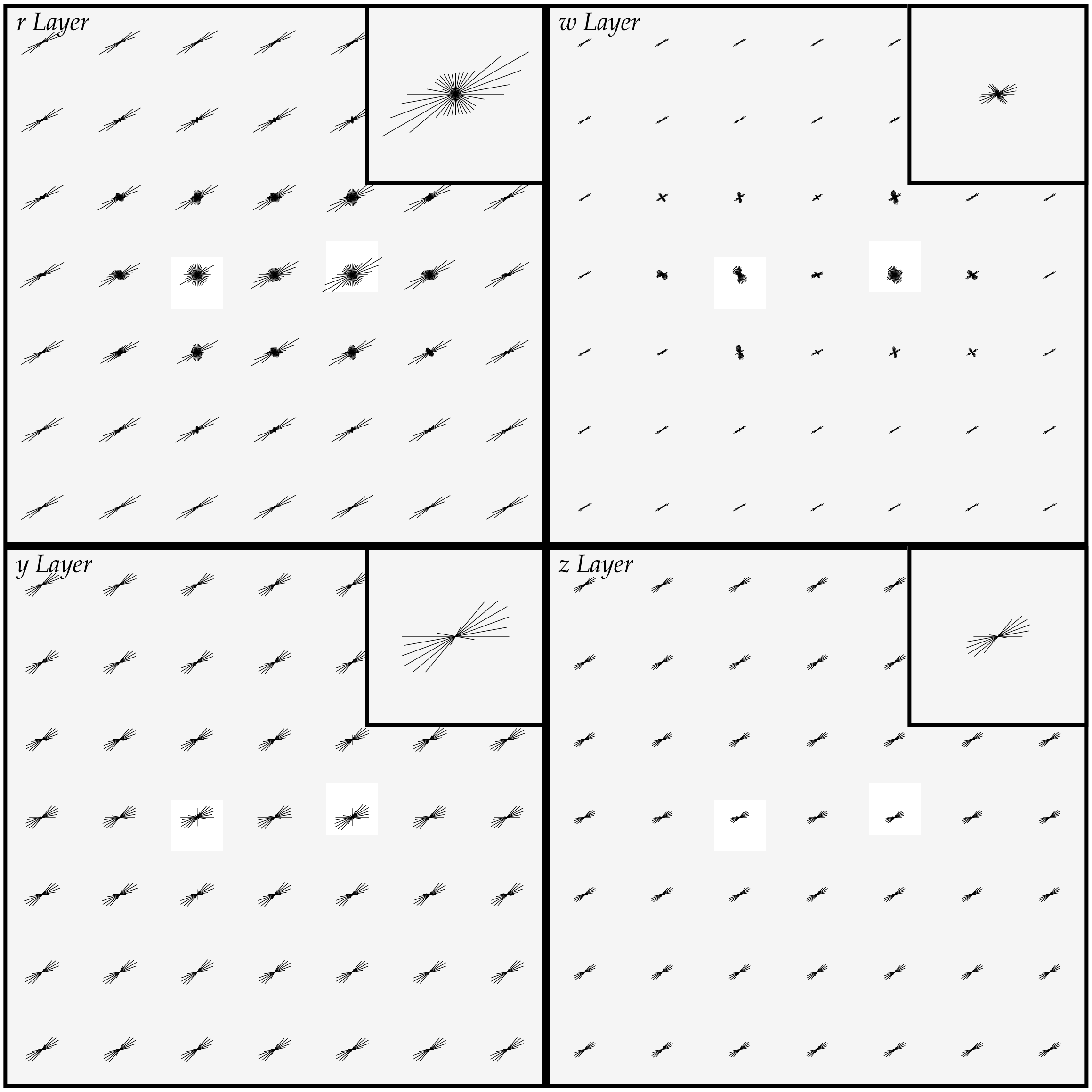


Figure 5-13: Needle plots of the layers in response to same polarity dots with a small gap on a +30° sinusoidal grating background. These plots show a 20 x 20 arcmin view, with an inset at the top tight showing a view of the center node. The sinusoidal grating dominates the responses of all the signals in all the layers. The signal of the dots is barely noticed and is seen as a distortion of the background pattern.

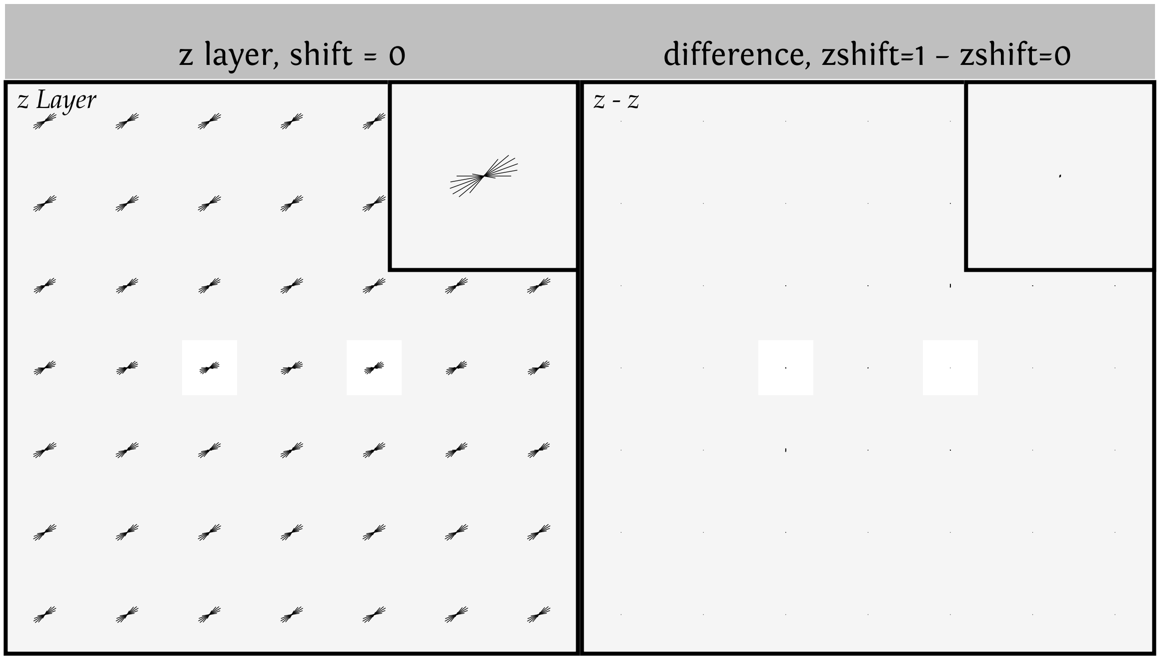


Figure 5-14: These plots show the same condition as the previous figure (Figure 5-13), a small-gap, same polarity stimulus on a +30° sinusoidal grating background. These plots show a 15 x 15 arcmin view, including an inset of the center node. The left plot (z) shows the z-layers for a shift of -1 (notice the very faint white squares showing the input). The plot on the right (z - z) shows the difference between the z layer with a shift of +1 arcmin and the z layer when the shift is -1 arcmin. Clearly, the differences supporting the detection in this case are very small.

5.3.6 Levi & Waugh Figure 10: Threshold Elevation

Now we can compare simulated results with actual data for the cases where there are sinusoidal grating backgrounds. The comparison with Figure 10 from (Levi & Waugh, 1996) is shown in Figure 5-15. The plot on the left shows data from the experimental paper and the plot on the right shows data from our simulations using the same conditions. As with Figure 4 from Levi & Waugh, there appear to be three separate regimes that have to be explained; the regimes are the same even though they are manifested differently, but are actually even more distinct in this representation.

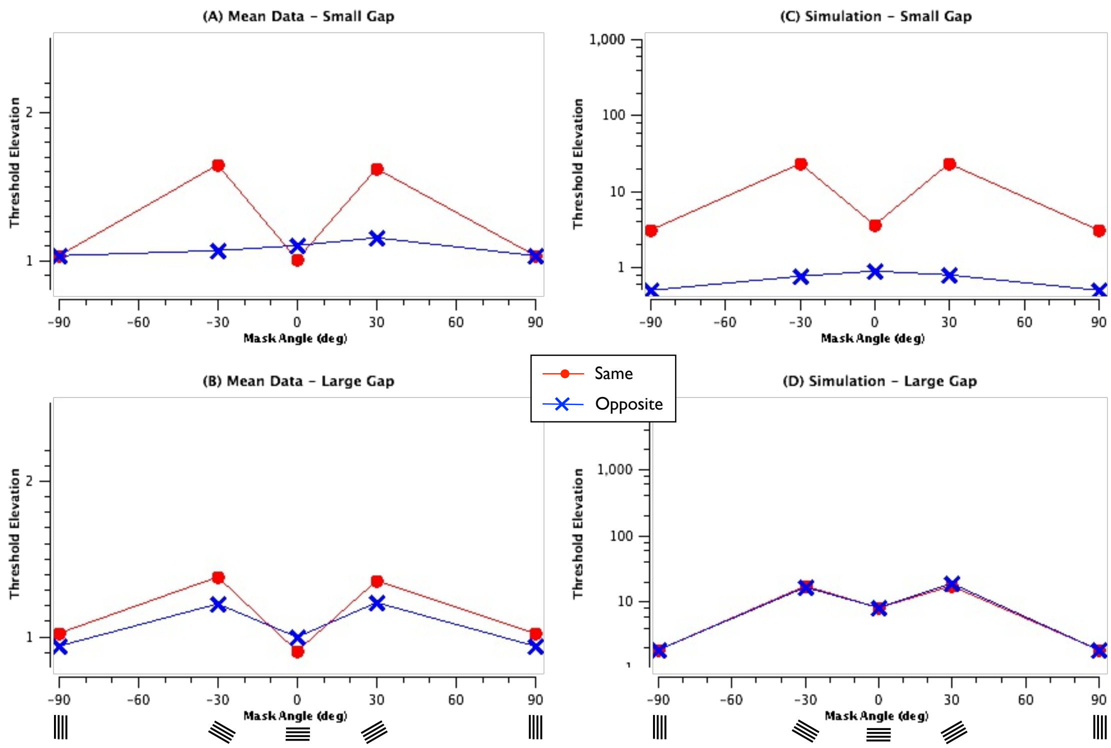


Figure 5-15: Threshold elevation for different orientations of a grating mask. Again, blue lines (crosses) denote opposite contrast stimuli and red lines (circles) denote same contrast stimuli. In the case of small gaps, the pattern for same contrast stimuli is quite different from opposite contrast stimuli. However, in the case of large gaps, there is no difference between the same and opposite contrast patterns.

Looking at the plots, there are clearly three different patterns of responses: small gap same-polarity, small gap opposite-polarity, and large gap for either contrast polarity. Each of these patterns is different, which suggests that the processing performed by the vision model is somehow different and distinct in each of these cases. The small gap same polarity case is susceptible to masking at 30° because it is operating at maximum sensitivity (e.g. hyperacuity) and the nodes that are used to measure the sensitivity are precisely at 30°. The small gap opposite polarity case is already being “masked” by the bimodal distribution of the filter responses, thus it is not affected by the grating. The large gap case relies on the long-range grouping mechanism to represent the shift of the dots; and while this mechanism can be masked, it does not start off being as precise as the pure filters in the small gap, same polarity case. Perhaps the filters nodes are more easily masked than the nodes representing long-range completion. The simulated data has the same patterns as the psychophysically measured data, indicating that the model represents real mechanisms.

5.3.7 Model Equivalent Psychometric Functions

When enough data has been collected in a psychophysical experiment it is possible to plot the observer response to the independent variable, the psychometric function. Since in this model we have a measure of the response before the decision is made, we can plot this response as the shift is varied. Plotting the model analog of the psychometric functions for several conditions on the same plot results in Figure 5-16.

It is interesting to note that the different slopes of the curves are due to the different sensitivities of the model for different conditions. The sigmoid shape of the functions is not due to any Gaussian noise, as there is no source of noise in the model, but rather to the design of the decision model. For stimuli with shifts close to zero, small differences yield curves that pass through the origin with a steep slope. As shifts move away from zero, the decision model weighted sum saturates and naturally turns into a sigmoid. The nonlinear output function accentuates this effect.

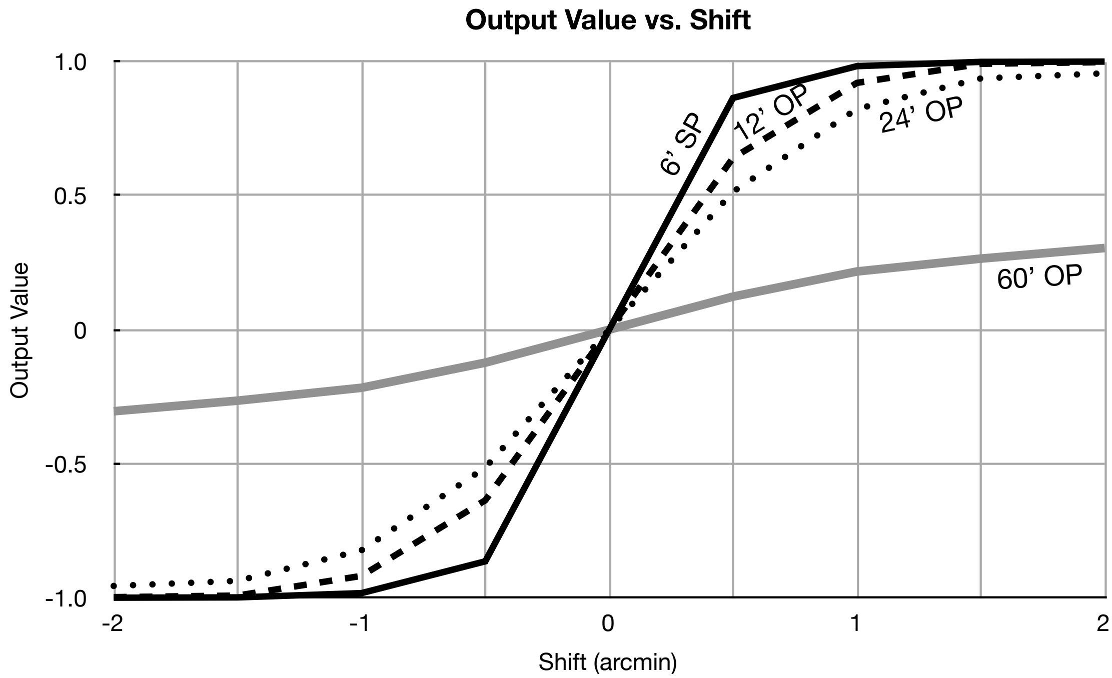


Figure 5-16: A demonstration of how different conditions can generate different response curves. The slope is different for the different conditions, thus the thresholds are different; the steeper slope indicates a smaller threshold. The conditions shown are (in order of steepest to shallowest): 6' gap same polarity (SP), 12' gap opposite polarity (OP), 24' gap OP, and 60' gap OP.

5.4 Discussion

The overall goal of demonstrating that a hierarchical neural vision model can indeed explain some of the more challenging data related to hyperacuity was successfully achieved. At the same time, some methods were tried that did not lead to success. There is as much to be learned from those approaches that do work as from those that did not, thus both will be discussed below.

5.4.1 Conclusions

One of the key questions at the outset concerned the so-called "readout" phase of the vision model. It was not assumed at the beginning of the project that the logical output layer would end up being the preferred readout layer. Thus one important question that needed to be answered in the context of these simulations is which layer (or layers), of the model are read-out, or output to decision making areas. One is inclined to think that the sharpened long-range completion layer provides the output; but it is possible that other layers provide types of information that is necessary for some types of decisions. In addition, if there is feedback within the model, then any of the layers within the feedback loop might also contain the required information. Surprisingly, it did turn out to be the case that the sharpened long-range completion layer provides the output. This was the conclusion after many different variations were tried, leading to a number of dead ends.

Another key question at the outset was the spatial location, perhaps moreso the spatial extent of the readout from the vision model. How large is the area represented by the vision model that provides input to the decision model? Again, after many trials and errors involving a rather large number of less than satisfying results, it was found that the location in the very middle of the visual field is sufficient to yield desirable results. It is reassuring that in both of these cases, (readout layer, and spatial extent of readout), that the common-sense a priori supposition actually makes sense and works.

5.4.2 Vision Model Issues

The vision model that we started with had to be adjusted in several way in order to produce the correct patterns of results. The particular changes made were: adding a second filter scale, combining the competitive kernels, post-processing (sharpening) of the boundary completion signals, and not including the feedback (from layer 5 to layer 3).

The initial formulation of the model had only one filter scale. A second, larger, filter scale was added in order to remove over-sensitivity from the model. The second scale adds robustness and allows the model to produce partially useful thresholds even when the most sensitive, smaller, scale is unable to produce reliable thresholds due to masking effects.

The competitive kernels were combined. Initially, the spatial competition preceded the oriented competition. This did not work, as it was found that the oriented competition must act on the filters before, or at least at the same time as spatial competition. If spatial competition alone acts on the filter outputs, then the filter outputs can no longer be used for threshold determination.

The feedback structure of the original model is kept in the description, even though it is not used. It seems that the feedback signal can be very useful for boundary completion of boundaries with large gaps as originally envisioned (Grossberg & Mingolla, 1985a, 1985b), thus it was not removed.

It was also found that a layer of sharpening beyond the boundary completion is useful. While perhaps not required, in the sense of the other modifications, it is still useful for producing a cleaner output.

On the other hand, it was found that the key model components of filters, spatial and oriented competition, and long-range completion were quite necessary. This is reaffirms the choice of model and is also supportive of more recent formulations of the Boundary Contour System (Grossberg, 2003, 2007; Grossberg & Yazdanbakhsh, 2005).

5.4.3 Decision Model Issues

As was the case for the vision model, the decision model also had to be adjusted in several ways in order to produce the desired pattern of results. In particular, it was necessary to adjust the learning, and think carefully about the use of feedback, that is, feedback in the sense of providing the observer information about the accuracy of responses given.

It was found that any sort of attempt to learn for each condition was met with failure. This produced what can best be described as "over-fitting", and is a tendency of certain neural networks that learn to try to accurately predict every piece of information they encounter. Instead learning of weights was done once before the individual conditions were presented, ad learning for each condition was restricted to the bias term.

The necessity for the use of response feedback was not immediately recognized. Feedback was provided to the human observers in the corresponding psychophysical tasks, and thus probably reduced some uncertainty in favor of a bias for each condition. While the feedback in the end was not used for learning the weights, it was still necessary, and was used to learn the bias term for each condition. In the original psychophysical experiments, it was noted that "The provision of feedback helped to overcome perceptual biases generated by the interaction between the mask and the target lines." (Waugh *et al.*, 1993), and thus it should be expected that the model would require the same information in order to match the human performance.

Simulations with Hebbian learning applied to each condition show that in some cases the response difference between two conditions *decreased* after learning was applied. A response difference decrease translates into a threshold *increase*. This is highly counterintuitive, and not simply related to the bias of the condition. Conditions with high bias (±30°) improved with learning just fine. The conditions that do not improve were very long-range (60 arcmin gap) and 0° grating (only the opposite polarity condition had a slight bias).

These results, together with the overall improvement when learning was performed once suggests that learning is not performed on any type of condition where there is any uncertainty about the correct answer. Instead, learning is only performed when choosing the correct answer is easy. To put it in another way, there is no learning on category boundaries.

5.4.4 Future Directions

There are a few issues that would lend themselves naturally to further investigation. It would be fairly straightforward to apply the model to a number of new stimulus types. It would be most interesting to find stimuli that cause the model to break down.

Also of interest would be to study jitter in the sampling structure of the model. As described, the model can easily support jitter of every sampling location for every layer, and this might be one contribution to noise in the model. One would expect basically the same pattern of results already achieved, but with some variability measures attached. It is not clear, *a priori*, that jitter would *add* noise; it might in fact produce thresholds with less noise (Yellott Jr, 1983).

It would also be useful to change the decision model to go from a single neural choice computation to use something like a neural antagonistic accumulator (Vickers, 1970, 1979; Smith & Vickers, 1988). This would allow for more than two choices, which have been used in some experiments (for example three choices: "left", "aligned", or "right"; or five position alternatives: "down two steps", "down one step", "aligned", "up one step", and "up two steps"). The neural antagonistic accumulator would also be a more neurally plausible mechanism as it would not need neurons that respond as well to negative signals as they do to positive signals.

Chapter 6: Conclusion

6.1 Summary of Work

The work reported in this dissertation concerns two separate but related aspects of hyperacuity research: an experimental effort, and a modeling effort. Both efforts have the goal of providing a better understanding of spatial vision models, especially Boundary Contour System (BCS) models, and help to limit the parameters of BCS models.

6.1.1 Experiments

The experiments performed were extensions of the Badcock & Westheimer (1985a, 1985b) paradigm that found a central zone of attraction and a surround zone of inhibition. The main results of the previous study were replicated. The full extent of the surround zone was mapped, and some new stimulus configurations were also introduced.

The original experiments used one configuration where the flanking line was split into two separate parts. Based on the work of Grossberg & Mingolla (1985a, 1985b), one might suppose that if only one of the split flanks were used, the repulsion in the surround would disappear, or at least be greatly diminished. This was tried, and there was no effect, as there is repulsion whether one or both of the split flanks are used. Although linking between the two flanks was not demonstrated, it cannot be ruled out either.

In addition, the effect of illusory lines was tested. They clearly cause repulsion in the surround zone, and there are hints that there is attraction in the central zone. The repulsion may be caused by the "illusory lines" formed by the co-linear "line-ends", or it may simply be due to the presence of "line-ends" in the inhibitory zone.

6.1.2 Simulations

Very small (hyperacute) stimulus differences were successfully simulated. The simulations performed were able to replicate the two chosen figures from Levi & Waugh (1996) that include several complex stimulus elements. In order to successfully recreate the psychophysically observed data, some modifications had to be made to the Boundary Contour System (BCS), and a decision model had to be included as well.

The main modification required of the BCS was combining the two competitive stages, the first competitive stage (spatial) and the second competitive stage (oriented), into one competitive stage with combined spatial and oriented kernels. The other observation is that the feedback is not required to explain the data, the feedforward circuit alone is sufficient.

It was necessary to include a modified Hebbian decision module into the model. By adding the decision module, it is possible to compute thresholds in a manner analogous to the psychophysical experiments that are simulated. In addition, the modified Hebbian decision module is able to include learning in two different ways that are essential.

With these modifications to the BCS and adding a decision module, the data were successfully replicated.

6.1.3 Finding Parameters

Both the experiments and the simulations were helpful in the overall goal of better understanding the BCS. The experiments helped map out the zone of inhibition, mediated by the (combined) competitive stage. The simulations provided even more insight into the structure and parameters of the BCS. The best set of parameters, used for all the simulation in Chapter 5, is listed in Appendix B (page 143) and should be useful for any future BCS simulations.

6.1.4 Explaining Hyperacuity

Using the BCS as an explanation for hyperacuity also makes it easier to see a resolution to the *filter* vs. *local-sign* theory debate that still lingers. The reason for the delayed resolution of this issue is that while many hyperacuity phenomena can be well explained by a filter model, there are some that remain unexplained, and how can they be explained except by using local-signs? Westheimer's (1981) modular local-sign hypothesis attempts to resolve this dichotomy, but is too vague to provide a satisfactorily explanation.

The BCS helps by showing that it is not just *filters* vs local-signs, but filters *plus other non-linear operations* such as center-surround competition and long-range grouping. These operations are "filters" as well, but they do add to the simple filter hypothesis and perhaps eliminate the need to consider local-sign explanations in the future.

It should also be noted that the proposed model using the BCS and a modified Hebbian decision module is the first biologically plausible model of hyperacuity. While previous models have taken some inspiration from biology, this model is based on neural networks with only local operations.

6.2 Lessons Learned

In this project, as in any serious undertaking, many paths are followed, some of which lead nowhere, others turn out to be very fruitful. By making note of some of these perhaps some effort can be saved by others in future efforts.

6.2.1 Hyperacuity

Simulating hyperacuity stimuli accurately is more complicated than was expected at first. In particular, the initial filter computation is extremely time-consuming because of the very dense sampling. As a result, these computations were stored in an intermediary database so that they did not have to be recomputed all the time. The other surprising aspect was the need for a decision model. For a long time different measures of model output were used before the need for a separate decision module was realized.

6.2.2 Learning in the Model

No learning happens in the BCS, the structure and parameters of the BCS stay fixed over the time-duration of the experiments. There are two kinds of learning needed in the decision model: learning of categories and learning of bias.

The learning of categories means learning whether a dot is shifted up or down or is level with the reference dot. This corresponds nicely with categorizing orientations, and in this work, this learning is performed before the experiment takes place. In fact, this learning does not change over the time-duration of the experiments. It appears that the learning of the categories ('up' vs 'down' etc.) only happens when the stimuli are unequivocally in one category or the other. When there is some uncertainty about the correct category, learning appears to be suspended.[11]

The other type of learning that takes place, and is much more malleable, is the learning of the bias term in the response of the decision model. This learning is relatively fast, and is used by the observer to learn the bias of particular stimulus configurations, especially when there is a grating of 30°. Because of the use of feedback in the original experiments (Waugh *et al.*, 1993), feedback is also used in the model to learn this bias term; the bias term is also learned for each stimulus configuration, so if there is a significant change in the stimulus, a new term will be learned.

While much attention has been paid to the transfer of learning in hyperacuity experiments (Sotiropoulos *et al.*, 2011; Weiss *et al.*, 1993), this was not necessary in the current case. The various stimulus conditions used were similar enough not to cause a problem.

6.2.3 Fitting Parameters to Complex Models

When fitting the parameters of a complex system like the BCS there is a tendency to rely on optimization algorithms such as Genetic Algorithms (GAs), this should be avoided as it is unlikely to help. In the BCS, problems are caused by the multiple layers and the strong interdependencies between parameters

---

11. This observation leads directly to a new ART variant. Perhaps the learning in ART should also be conditioned on the certainty of the match. As traditionally implemented, when a category is chosen and the top-down pattern matches better than the vigilance threshold, learning is allowed to take place. Perhaps and additional constraint should be placed on the learning, such that if the bottom-up choice is 'close' somehow, then learning is suspended. Then learning along competitive boundaries would be slower. This may already happen in ART-3 networks, but it should be incorporated and studied in an algorithmic version as well.

6.2.4 Boundary Contour System (BCS)

The BCS provided the most issues to deal with. As mentioned previously, the spatial and oriented competitive stages have been combined into one, it was found that the feedback is not necessary, but that another stage of sharpening is necessary at the output stage.

One of the questions that was unclear at the outset was what stage would be the actual output that would be used by the decision model. It is not clear a priori that the output of the long-range completion is the only alternative to consider. In fact, many options were tried, but in the end, the output of the long-range completion, passed through a sharpening stage is exactly what made the most sense. Before this decision was finalized, other possibilities were considered, even the simultaneous output of many different stages to the decision model. This would then allow the decision model to choose the level that would be most appropriate, which might be different for different tasks. However, there was no need for direct filter learning, as in the reverse hierarchy theory (Ahissar & Hochstein, 2004) The output of long-range completion is sufficient.

6.3 Remaining Issues and Future Work

6.3.1 Decision Module Using Accumulator Model

The decision model used in this effort is one of the simplest possible. It can make a binary decision, a yes/no, or a up/down distinction. For a more general decision module that can make more than two choices a more complex design is needed. There have been suggestions that the accumulator design is more physiologically accurate (Smith & Vickers, 1988; Vickers, 1970, 1979; Vickers &

Smith, 1985; Vickers & Lee, 1998, 2000), and a design along that line of thinking should be explored. Such a model would likely also provide better estimates of actual reaction times. In any case, it is not expected that the move to such a model would alter any of the results presented in this dissertation.

6.3.2 Model the Effects of Jittered Sampling

In a similar vein, the way the model is conceptualized, described, and expressed in the code for the simulations makes it possible to use a jittered sampling grid. It may even be possible to completely move away from the concept of grid and just provide locations that need to be sampled in order to compute the next layer in the model. The sampling between different layers need not use the same grid. It is unclear what properties this type of sampling would add to the model, it is possible that it would add noise, reduce noise (Yellott Jr, 1983), increase overall accuracy, decrease overall accuracy, etc. It is not expected that modifying the sampling in this fashion would change any of the results presented in this dissertation.

6.3.3 Explain More Hyperacuity data

There are several hyperacuity displays that are still challenging. Some involve moving stimuli, for example the hyperacuity observed in displays defined by moving stimuli (Westheimer & McKee, 1977a). This is tricky because any explanation would probably have to rely on some theory of moving stimuli. There is one possible explanation that would fit nicely with the proposed model, and it relies on the fact that the initial filter outputs are summed together irrespective of spatial phase. If the summation of the moving hyperacuity displays (Westheimer & McKee, 1977a) is indeed within 4 arcmin, then the separate displays can all be represented by the same cells in the rectified layer (roughly analogous to *complex* cells in V1). The width of these rectified cells at the highest spatial scale is about 4 arcmin. Thus the model should be able to handle these moving stimuli with only some tweaking to deal with very brief stimulus presentations, and summation over *time* in the rectified layer. Everything else would be unchanged.

The experiments of Badcock & Westheimer (1985a, 1985b) and also those presented here in Chapter 3, might also need some type of motion theory in order to properly simulate the various conditions, since the displays change over time. It may be possible to somehow avoid this by comparing static displays; but given the essential motion quality of the stimuli, it seems unwise to completely ignore this aspect of those experiments. Again, perhaps a simple modification to the rectified layer to deal with summation over time would suffice.

Some challenging hyperacuity data do not involve moving displays, for example the data of Watt & Campbell (1985) may also be difficult to account for. They displayed Vernier stimuli in random orientations close to the vertical and found that thresholds were much higher than when stimuli were displayed in a fixed vertical orientation. One possible way to explain this data is that the visual system may be highly attuned to orientations that are exactly vertical and horizontal. Westheimer (2005) found such effects in the periphery.

Another interesting set of stimuli are those of Morgan *et al.* (1990), which interpose distractor dots between the dots of a Vernier display. The model should be able to deal with those displays with minimal changes. The distractor dots would likely add noise to the results of the model, increasing the thresholds slightly in all conditions, but more severely when the distractors are near the target dots. Furthermore, it is likely the systematic effects of bias can be predicted when the distractor dots are in specific locations, thus shifting the bias of the psychophysical function, not just increasing the observed thresholds.

6.3.4 Explain Visual Illusions

Many other visual stimuli, especially visual illusions, that have not yet been explained with the BCS, should now be within reach. It is very likely that the types of spatial processing which has been investigated are closely related to various metrical and angular illusions such as the Müller-Lyer illusion (with its multitude of variations), Poggendorf's illusion, Herring's illusion, Wundt's illusion, Zöllner's illusion, just to name a few. With a reasonable set of parameters for the model, it might be worthwhile to simulate some of these illusions; since the illusions may be explainable using the same mechanisms as the hyperacuity tasks. In addition to the visual illusions mentioned, there is also the data on orientation contrast provided by Nothdurft (1992). This data might also be worthwhile to simulate with the model. A related approach, but focusing on texture boundaries, is explored by Bhatt *et al.* (2007)

6.3.5 Transfer of Learning

Poggio & Fahle (1992; Fahle *et al.*, 1995) have achieved some success by using a HyperBF model to study transfer of learning from one hyperacuity paradigm to another. It is possible that transfer of learning may be a straightforward application of the BCS presented here as well.

By transfer of learning is meant that once the subject has become proficient at performing one hyperacuity task, this may help the subject in a slightly different hyperacuity configuration. Tasks which are related and therefore should show transfer are: orientation discrimination (Westheimer, 1981), two-dot Vernier, and Vernier with flanking lines. On the other hand, there should be no transfer of learning for stimuli of different orientations; indeed if the orientations are close, there may be significant interference! This type of effect may be relevant to the Watt & Campbell (1985) study discussed previously.

6.3.6 Data Bootstrapping

Finally, an intriguing idea that would "bootstrap" the model in part would be to simulate the experiments of Wilson by which he derived the shape of the psychophysical masking function (Wilson & Gelb, 1984). The competitive and cooperative interactions of the model might distort the receptive fields of the initial nodes; for example, it is possible that the initial filter bank consists of Gabor filters, and when probed with the Wilson masking technique, a different shape might be seen!

Appendix A:
Phenomenology of Experiments

This appendix contains some interesting observations relating to the phenomenology of the experiments described in Chapter 3 of this dissertation (see Figure 3-6 on page 42 and Figure 3-7 on page 44). The experiments have a very simple structure; that is, they all consist of two static images displayed sequentially. Despite this apparent simplicity, they often induce a percept of motion. Furthermore, this percept of motion varies in a number of ways with non-obvious changes in the display parameters. Because these effect relate to motion perception, they are not relevant to the analysis of static displays which concerns most of the dissertation, especially the modeling effort.

This is an attempt to catalogue some of the different percepts, with the hope that someday more quantitative relations can be established. In some ways the changes are reminiscent of the Ternus effect, although not as easy to demarcate. Since the Ternus effect has had a positive impact on modeling, careful study of these effects might also lead to fruitful modeling efforts. After all, it is not satisfactory to leave the explanation of these effects to an unknown higher-order mechanism. It is also necessary to mention that it is highly unlikely that the differences in percepts described can be attributed to variability of attention, since the experimental paradigm focuses attention on the target line in all instances.

What kinds of percepts are seen in these experiments?

(1) There is almost always a perception of motion. Sometimes the motion is very definitive and clear. At other times it is vague, and one only has an uncertain sense of motion. It is likely that the motion is more vague when the displacement is smaller.

(2) When the target line and flanking line are close (perhaps ~3 arcmin), it often appears that the lines are moving away from each other.

(3) At other times, it often seems that the flanking line is moving very quickly past the target, in order to stop where it is. This type of movement of the flanking line is perplexing, since the flanking line only appears in the second display.

(4) The motion of the target is sometimes a short, instantaneous jump.

(5) Sometimes the target and the flanking line appear to move *together*. This means that the flanking line is first perceived, and then both move. This is of course not in accord with the actual sequence of events since the flanking line only appears in one location, and at the same time as the target moves (actually, the target has already moved). This is one of the more bizarre perceptions.

(6) There are times when the flanking line seems to emerge from the target line!

(7) At other times the flanking line seems to emerge from the surround (recall that the background is framed by a window 2° x 2°). This is especially true when the flanking line is far away from the target line (~30 arcmin). Perhaps the sole determinant is which source of brightness is closer.

(8) And of course, there are times when the flanking line emerges out of nowhere; in other words, it just pops up.

As for possible explanations of these observations, the Grossberg & Rudd (1992) explanation of the Ternus effect involves interactions between low- and medium-level motion processes, it is quite possible that the phenomena described (especially #5) can be adequately described using similar modeling techniques. Another possibility is that there are interactions between the processing of static images and motion processing.

Appendix B:
BCS Variables & Parameters

The main purpose of this appendix is to present the values of the parameters used in the simulations, these are shown in Table B-1. The table shows each parameter, the symbol used, the values used in the simulations, and the equations where the parameter is used.

It was mentioned earlier that the simulation use initial filters that are Gabor (1946) shaped and based on the initial filters determined psychophysically by Wilson (1986). The parameters for the Wilson filters are shown in Table B-2.

Table B-1: Parameter values used in simulations. This table shows each parameter, the symbol used, the values used in the simulations, and the equations where the parameter is used.

Parameter	Symbol	Value	Comment / Equation
General			
Width of visual field	–	31	in columns
Height	–	21	in columns
Orientations	$180/\rho_k$	18	Equation 5-4
1/Scale	ρ_{pq}	4.5'	column spacing, Equation 5-4
Iterations	–	1	model always feedforward
filter sampling	–	0.09	sampling density of filters
falloff limit	–	0.01	extent of filter sampling
background	–	100	input background intensity
Filters – Scale 1			
filter frequency	ω	9.0	Equations 5-6 and 5-7
filter scaling factor	$1/\lambda$	600	implied in Equation 5-9
filter width	σ	4.0	Equations 5-6 and 5-7
filter aspect ratio	α	1.7	Equations 5-6 and 5-7
Filters – Scale 2			
filter frequency	ω	4.5	Equations 5-6 and 5-7
filter scaling factor	$1/\lambda$	150	implied in Equation 5-9
filter width	σ	8.0	Equations 5-6 and 5-7
filter aspect ratio	α	1.7	Equations 5-6 and 5-7
Rectification Stage – r			no parameters!
Feedback Stage – v			
excitatory input scaling	D_v	2.0	Equation 5-10 and #-#
bottom-up ratio	β	0.67	Equation 5-10 and #-#
output scaling	W_v	1.0	implied (see Equation 5-3)
Competitive Stage – w			
excitatory input scaling	D_w	5.0	Equations 5-13 and 5-14
lower bound	L_w	5.0	Equations 5-13 and 5-14

tonic activity	T_w	0.0	Equations 5-13 and 5-14
center kernel spatial width	$\sigma_{pq}^{c_w}$	0.5'	Equation 5-11
center kernel oriented width	$\sigma_k^{c_w}$	13°	Equation 5-11
surround kernel spatial width	$\sigma_{pq}^{s_w}$	8'	Equation 5-12
surround kernel oriented width	$\sigma_k^{c_w}$	33°	Equation 5-12
output scaling	W_w	1.0	implied (see Equation 5-3)
Long-range Grouping Stage – y			
each lobe max contribution	H	0.5	Equation 5-18, fixed value
lobe shunting parameter	K	0.6	Equation 5-18
bipole distance falloff	σ_ρ	25	Equation 5-15
bipole location falloff	σ_θ	20	Equation 5-15
bipole tangent falloff	σ_φ	20	Equation 5-15
spatial impenetrability threshold	τ_{si}	0.3877	Equation 5-19
output threshold	τ_y	0.5	Equation 5-20
output scaling	W_y	2.0	Equation 5-20
Sharpening Stage – z			
excitatory input scaling	D_z	6.0	Equations 5-23 and 5-24
lower bound	L_z	5.0	Equations 5-23 and 5-24
center kernel spatial width	$\sigma_{pq}^{c_z}$	0.5'	Equation 5-21
center kernel oriented width	$\sigma_k^{c_z}$	20°	Equation 5-21
surround kernel spatial width	$\sigma_{pq}^{s_z}$	0.5'	Equation 5-22
surround kernel oriented width	$\sigma_k^{c_z}$	50°	Equation 5-22
output scaling	W_z	1.0	implied (see Equation 5-3)
Decision Model			
learning parameter	ϵ	0.001	Equation 5-27
response nonlinearity	γ	5.0	Equation 5-26
learning iterations	iterations	100,000	

Table B-2: Filter Parameters. Gabor parameters matching the Wilson parameters. See Wilson (1986) for equations and explanations.

Parameter	Scale 1	Scale 2	Scale 3	Scale 4	Scale 5	Scale 6
Peak (cpd)	0.8	1.7	2.8	4.0	8.0	16.0
A (–)	30.0	70.0	140.0	150.0	76.7	18.4
B (–)	0.267	0.333	0.894	0.894	1.266	1.266
C (–)	–	–	0.333	0.333	0.500	0.500
σ_1 (deg)	0.198	0.098	0.084	0.059	0.038	0.019
σ_2 (deg)	0.593	0.294	0.189	0.132	0.060	0.030
σ_3 (deg)	–	–	0.253	0.177	0.076	0.038
σ_4 (σ_1)	3.2	3.2	3.2	3.2	3.2	3.2
H (–)	1.097	0.950	0.700	0.760	1.343	1.400
ε (–)	0.55	0.55	0.55	0.55	0.25	0.25
Spacing (σ_1)	0.56	0.56	0.56	0.56	0.56	0.56
Spacing (°)	30.0	30.0	20.0	20.0	15.0	15.0
in cycles	0.16	0.17	0.24	0.24	0.30	0.30
L	21.99	46.69	61.46	65.85	17.948	4.306
w	0.01671	0.03556	0.05235	0.07464	0.143	0.285
s	22.324	12.383	10.803	7.534	4.400	2.186
a	1.703	1.519	1.493	1.504	1.658	1.669
Actual peak	1.00	2.13	3.14	4.48	8.58	17.1
1/4 cycle	15.0	7.03	4.78	3.35	1.75	0.877
1/4 cycle	1.26	1.20	0.95	0.95	0.77	0.77

Appendix C: Psychophysical Methods

When an experimenter desires to know the response function of an observer for different values of a stimulus parameter, this is called the psychometric function. There are several different types of psychometric functions, each appropriate for different types of experiments. This appendix considers experimental paradigms, psychophysical methods, and other issues related to psychophysical functions.

C.1 Experimental Paradigms

Two frequently used experimental paradigms are two-alternative forced choice (2AFC) and Yes-No. In both cases the observer must respond with one of two responses on each trial. The main difference is that in 2AFC the observer is presented with two stimuli, and must discriminate between signal and noise. The 2AFC paradigm has been found to reduce bias in responses (Macmillan & Creelman, 1990) and is especially well suited to detection measurements. In Yes-No experiments only one interval is presented, and the observer decides whether the signal is present or not. The Yes-No paradigm can also be used to discriminate between two events, the analysis which applies is identical to the signal/noise case however.[12]

12. It is not clear that there is an experiment which can be performed in one paradigm but not the other; for example, one of the events can be classified as signal, whereas anything else is noise. Efficiency and the straightforwardness of the task should be the main considera-

Thus the question answered by the observer in the 2AFC case is: In which interval was the signal presented? The interval here can be spatial or temporal; that is, the two alternatives can be presented simultaneously or in serial order. The question answered by the observer in the Yes-No paradigm is: "Was the signal present?" or "Did event A or event B occur?"

In any case, the psychometric functions for these two cases are different. The shape in either case is an ogive, but the 2AFC curve starts at 50% whereas the Yes-No curve starts at 0%. Examples of these curves are shown in figure C–1. Clearly, the 2AFC case can be generalized to more choices such as 3AFC and 4AFC. In these cases the curve will instead start at 33% and 25% respectively.

---

tions when choosing a paradigm; proper application of signal detection theory will yield the same results for either paradigm. 2AFC is very efficient when both alternatives can be presented simultaneously. However, if the stimuli need to be foveated or otherwise carefully attended it becomes necessary to present them serially. With serial presentation of the two alternatives, the paradigm is really no different from the Yes-No paradigm. In most cases where serial presentation is required it is more efficient to collect a response after each stimulus presentation (Yes-No) rather than having the observer respond after every other stimulus (2AFC).

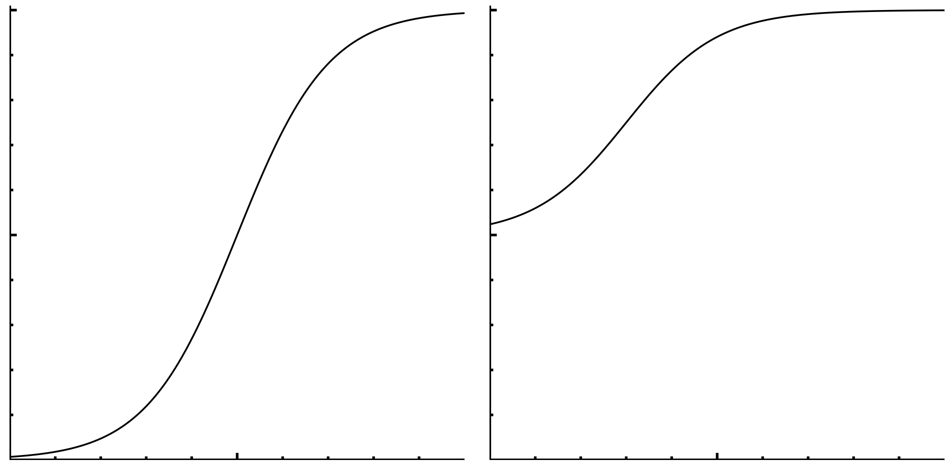


Figure C-1: Psychometric functions for different experimental paradigms. On the left, a Yes-No paradigm will yield probabilities from zero to 100%. On the right, a 2AFC starts with a minimum of 50% and maxes out at 100%.

If the psychometric function is estimated for a sufficient number of stimulus levels, it will clearly resemble an ogive curve when plotted. The precise mathematical form of this function is not agreed upon. Signal detection theory (SDT) predicts that the shape is the cumulative distribution of the underlying noise distribution. SDT usually assumes that the noise has a Gaussian distribution, and therefore the psychometric function is often assumed to have a cumulative Gaussian distribution. Choice Theory (Luce, 1959, 1963b, 1963a) has a different theoretical foundation which predicts that the psychometric function has the shape of a logistic function.[13] The logistic function is a simple mathematical function whereas the the cumulative Gaussian must be computed using numerical approximations. In practice however, the shape of these two functions (logistic and cumulative Gaussian) are almost identical and thus it irrelevant

13. Link (1992) has presented a theory which also predicts that the shape of the psychometric function should be the logistic function. This theory is quite different from Luce's. It is rather interesting that two different deeply theoretical analyses converge on this point.

which one is used. There is a third function which is often used, the Weibull function. This function has nice theoretical properties and is often used where there is a *logarithmic* dependence on the stimulus parameter, such as intensity of light or sound.

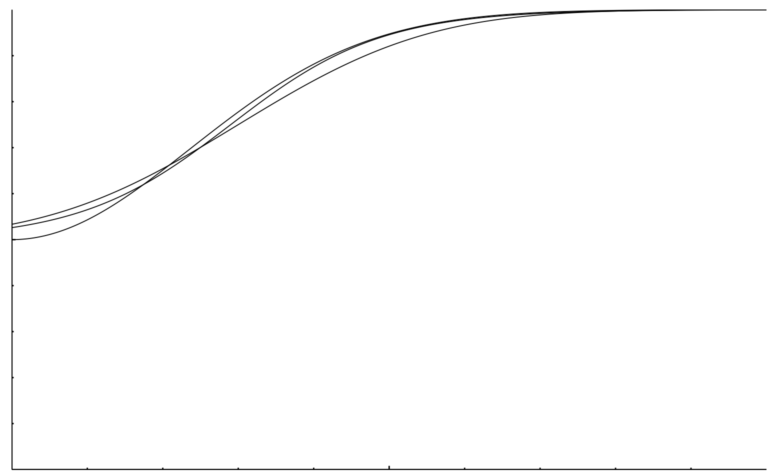


Figure C-2: The shapes of different psychometric functions. The logistic, cumulative Gaussian, and the Weibull function (order from top to bottom on the y axis) plotted for a 2AFC experiment. Clearly the functions are very similar, and can be substituted for each other in many cases.

The Yes-No paradigm was used for all the experiments presented in Chapter 3 of this dissertation. The question presented to the observers was: Did the line move to the right or to the left? This task is clearly an example of the Yes-No paradigm. And since it is a task which requires focused spatial attention stimuli would have to be presented serially in an 2AFC paradigm, rendering it less efficient.

The logistic function has been used as an approximation to the shape of the psychometric function. It seems to fit the data well, but the primary reason is the ease with which it can be computed.

C.2 Psychophysical Methods

The first and simplest method for finding the psychometric function is called the method of constant stimuli (MCS). The idea is to choose a number (usually about half a dozen) levels of the stimulus and then to present the observer with a number of trials at each of these levels, recording the observers responses. By combining the responses at each level, the probability of a particular response can be determined and plotted. A mathematical form of the psychometric function can even be fitted to the estimated probabilities using curve fitting methods. For example, to measure contrast sensitivity at a particular spatial frequency a number of contrast levels are chosen. These contrast levels are known to span the psychometric function. For an accurate determination of the psychophysical function, a total of 200–250 trials may have to be run, spread over the different contrast levels.

The MCS as stated works quite well, but it has a couple of problems. The first problem is that the location and shape of the psychometric function must be fairly well known before the experiment is started! This means that many hours are spent in pilot experiments determining the exact stimulus levels to be used in the experiment. The other problem is that the method is quite inefficient. Many trials are presented at levels which are unecessary and which contribute minimally to the determination of the psychometric function. There is a trade-off between these two problems of course: the more well-known the location and shape of the psychophysical function is, the fewer levels need to be measured, however this requires more time spent on pilot experiments determining the location and shape. If there is a lot of individual variation in the location of the psychometric function, then extensive pilot measurements do not save much time.

Because the MCS needs either a lot of pilot experiments or testing at many levels, other methods for finding the psychometric function have been developed. The ideal method requires minimal pilot experiments and will make the most effective use of observer trials.

A number of methods known as "staircase" procedures are very commonly used in psychophysical experiments. The basic idea is to start with a stimulus testing level which is known to produce a "Correct" (or "Yes") response. Then the following two step procedure is iterated: (1) Present a number of trials and record the responses. (2) If the responses are all "Correct", then take a step to a more difficult stimulus testing level, if one or more responses is "Wrong", then shift to an easier stimulus testing level.

The various methods differ in how many trials are used at each stimulus testing level; normally 2 (for a 71% staircase) or 3 (for a 79% staircase) are used. There are also several algorithms for adjusting the stepsizes used in moving from one testing level to another. The common characteristic is that the steps get smaller over time. Algorithms also differ in when they stop the iterative procedure. Some stop after a pre-determined fixed number of trials; others stop after a certain number of "reversals", that is changes from stepping in the easier to the harder direction or vice versa. An algorithm with adjustable step size can stop when the steps become smaller than a certain tolerance.

The result of most staircase algorithms is the stimulus testing level at which the staircase converges. A 71% staircase converges to the stimulus level where the subject will respond correctly 71% of the time. It is possible to use the final testing level as the result; more common is perhaps averaging the testing levels at the reversal points. If all the testing levels and the responses are saved, then a curve fitting routine can be applied to all the data collected. This is often a better way to estimate the psychometric function as it is possible to get error estimates for fitted parameters.

Probit analysis (Finney, 1952, 1971), the procedure commonly used to fit a psychometric function to data collected with MCS and other methods, merits a more thorough discussion. It is essentially a procedure for fitting a non-linear curve (an ogive) to a set of data points given that the computational power available is limited. Fitting a straight line to a set of data-points is simply done using the method of least-squares and can easily be done without the use of a computer. The idea behind probit analysis is to make a guess about the psychometric function (usually a cumulative Gaussian) and improve upon this guess iteratively. The steps to get an improved guess are as follows: (1) Compute z-scores for the probabilities given by the data at each stimulus level using the current estimate of the psychometric function. (2) Fit a straight line to the plot of z-score vs stimulus levels. (3) Convert the parameters defining the line back to parameters defining the psychometric function.

The use of z-scores implies that only cumulative Gaussians can be fitted, but this is not the case. If a logistic functions is to be fitted, then the probabilities have to be converted to the equivalent of z-scores for logistic functions. Other than this modification, the procedure is exactly the same as when using cumulative Gaussians. Clearly, this procedure allows the use of any monotonic function. One very useful feature of probit analysis is that it provides estimates of the error of the parameters estimated.

There is a more general way to provide error estimates of the parameters estimated in a particular procedure which is irrelevant of the actual procedure used. This method is known as the Monte-Carlo method. The method is illustrated in Figure C-3. The best parameters of the model are estimated from the data collected. These parameters are used in the model to generate a number of synthetic data sets, each of which is used to estimate a set of parameters. The distribution of these parameter sets is used as the distribution around the "true" parameters which will remain unknown. This is the main assumtion of all Monte-Carlo simulations: that the distribution of estimated parameters around the true parameters is the same as the distribution of parameters nearby in parameter space.

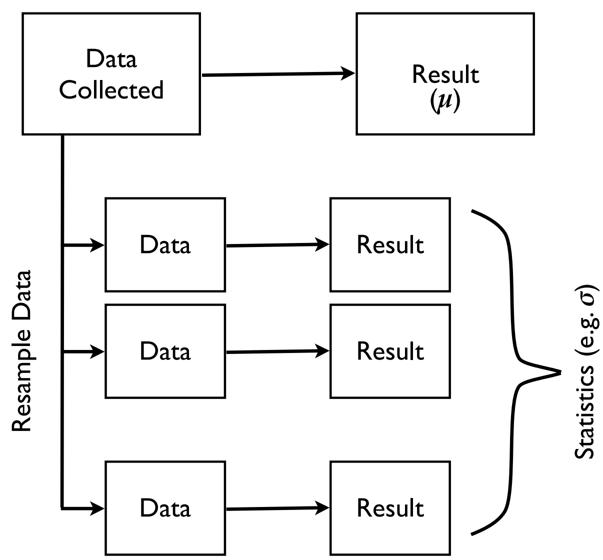


Figure C-3: The Monte-Carlo method. The complete data set is used to estimate the mean of the chosen parameters. The data set can be randomly resampled in order to generate statistics such as the standard deviation for each of the parameters.

The Monte-Carlo method for estimating errors can be used together with probit analysis. Foster & Bischof (1991) tested a Monte-Carlo method which is known as the Bootstrap to statisticians and found that it compared favorably with the standard method used with probit analysis.

Probit analysis relies on least-squares fitting, which assumes that data-values have a normal distribution (for a nice explanation see Chapter 14 in Press *et al.* (1986)). If the data is collected using the MCS, then there need to be enough trials at each level (> 30) to be able to compute reliable statistics. However, if the data is collected using a staircase procedure, there may be many fewer points at each level, and a binomial distribution should be used instead. A more severe problem are those levels where the measured probability is one of the extremes (0% or 100%). The conversion to z-scores places these at $\pm\infty$, therefore these probabilities have to be adjusted before they can be used. Probit analysis will not work if there are too many of these cases.

For these reasons, both theoretical and practical, we need a different procedure to fit a nonlinear function to binomial data. Such a method exists, and it is commonly called maximum likelihood estimation (MLE). There is some confusion here, because least-squares fitting is also a MLE method, but is not usually referred to as such. This convention will be adhered to.

The MLE follows directly from the simple mathematical desciption of the probability of the measured responses given the probability of response at each stimulus level, see Equation C-1. The description of the probability of response at each level is of course the psychometric function. Thus the total probability of the data can be written in terms of the parameters of a certain psychometric function, and is fairly easily computed. MLE is the idea of maximizing this probability by adjusting the parameters of the psychometric function. This has to be done using some numerical maximization algorithm. There are many to choose from and books have been written on this subject.[14]

$$P(r_1, r_2, ...r_n) = \prod_{i=1}^{n} P_{r_i}(s_i) \tag{C-1}$$

A couple of psychophysical methods have been developed based on MLE. The idea is to use the responses of all the trials to get a current estimate of the psychometric function. From this estimate it is possible to calculate what the stimulus testing level for the next trial should be in order to *maximize the efficiency* of the testing procedure. There are two MLE methods often used by experimenters, Best PEST (Pentland, 1980) and QUEST (Watson & Pelli, 1983). Both of these methods were developed in order to work efficiently on computers of limited computational power. Both of the methods estimate the location of the psychometric function and assume a fixed slope. QUEST uses an a priori estimate of the psychometric function to aid the process, but the prior information is excluded when computing the final result.

Both procedures use the idea of the minimum sweat factor (Taylor, 1971) when deciding what the next stimulus testing level should be. The idea is to reduce the variance of the measurement while also testing at a stimulus level which will provide maximum information about the location of the psychometric function.

The two MLE methods mentioned, Best PEST and QUEST, are very useful for finding the location of psychometric function when the slope is approximately

---

14. Again, see Press *et al.* (1988). Chapter 10 has a nice selection of maximization/minimization algorithms.

known. However, there are cases when the precise value of the *slope* is what one wants to know. In these cases, the MLE methods are not likely to be very efficient, since they are optimized for finding the *location* of the psychometric function. The MLE methods do work if the maximization algorithm can simultaneously adjust the location and the slope parameters of the psychometric function. The problem is that the data has been collected in such a way as to minimize the error of the estimate of the location parameter. The error of the slope parameter may be much larger. For this reason, neither of the MLE methods discussed are adequate for the experiments presented in this dissertation.

The Adaptive Probit Estimation (APE) method (Watt & Andrews, 1981) performs preliminary estimates of the psychometric function and uses these estimates to decide what values to use for the next set of trials (10). Each trial is randomly chosen from four possible values, two on each side of the midpoint of the preliminary psychometric function. Two points are near the middle (thus more difficult for the observer) and two are farther away from the midpoint (easier for the observer). The points are randomly chosen from this set, but the method also ensures balanced coverage. While the method efficiently homes in on the best range of values to use for a given subject for a given condition, it is difficult to display the collected data and its fit to a psychometric function, because data is collected for so many different points with only a small number of trials for each stimulus value. This problem and one suggested remedy, binning, are illustrated below.

A new procedure, Adaptive Logistic Estimation (ALE) is suggested that improves upon the the previous methods. It uses a logistic function instead of the cumulative Gaussian (with its probit calculations). Because the logistic function can be expressed as a simple calculation, the maximum likelihood estimation (MLE) of the parameters for the logistic version of the psychometric function can be done in a straightforward manner.

The MLE procedure merely finds the most probable logistic function from which all the data collected for a condition could have come. To calculate the error estimate, the Bootstrap method (a Monte-Carlo method) uses the psychometric function estimated from the data to generate one hundred simulated data sets. These data sets have the same number of measurements at the same stimulus levels as the collected data set, but with data generated by a pseudo-random function and the best-fitting psychometric function. Each of the simulated data sets are used to estimate a separate psychometric function. Then statistics are computed for the collection of simulated functions to provide error bounds for the estimated function. This method works because the statistics of the simulated psychometric functions are assumed to be closely related to the true statistics around the estimated function. This assumption is reasonable as there is no expectation that the distribution of errors would change appreciably for related psychometric functions (Press *et al.*, 1986).

One might hope that since both the logistic function and the maximum likelihood estimation formula for binary responses are nice analytic expressions it would be possible to solve analytically for slope and the location of the psychometric function. That is not the case.

C.3 Best Parameters of ALE

Some obvious questions present themselves when faced with a fairly complex adaptive procedure. Should the range of stimuli used be wide or narrow? How many levels should be chosen for each block? It is obvious that the more trials are used, the better the results; but perhaps the most important question is how many trials are really necessary?

These questions are easily answered by some simple simulations. The goal for choosing any of the ALE parameters is to reduce the standard error of the final estimate. Since the Bootstrap method will be used to find this standard error of actual data it can be used for these simulations as well. The idea is to apply the Bootstrap to various combinations of stimulus levels, ranges, and number of trials in order to determine which ones work the best.

In order to determine both the slope and the location accurately, it is sufficient to continually estimate 4 level, the range of the the levels should be about 4.4 times the spread (inverse of slope), this covers about 10%–90%. A total of about 250 trials is enough.

C.4 Displaying Raw Data Collected with ALE

It is a straightforward matter to display data collected with the MCS and to compare it with a fitted function of a specific shape. The trials for each stimulus level probed are combined to yield a probability estimate for that level. This simple procedure is inadequate for data collected from an adaptive routine. There are two difficulties with displaying the raw data in such a non-parametric manner: (1) the data is collected at very many stimulus levels and (2) a different number of trials may have been used at each level. Since this often means that only a few trials have been presented at any one level, the data points tend to cluster on the 0%, 50%, and 100% lines. If the data points are plotted this way, the plot becomes a total mess as shown in Figure C-4.

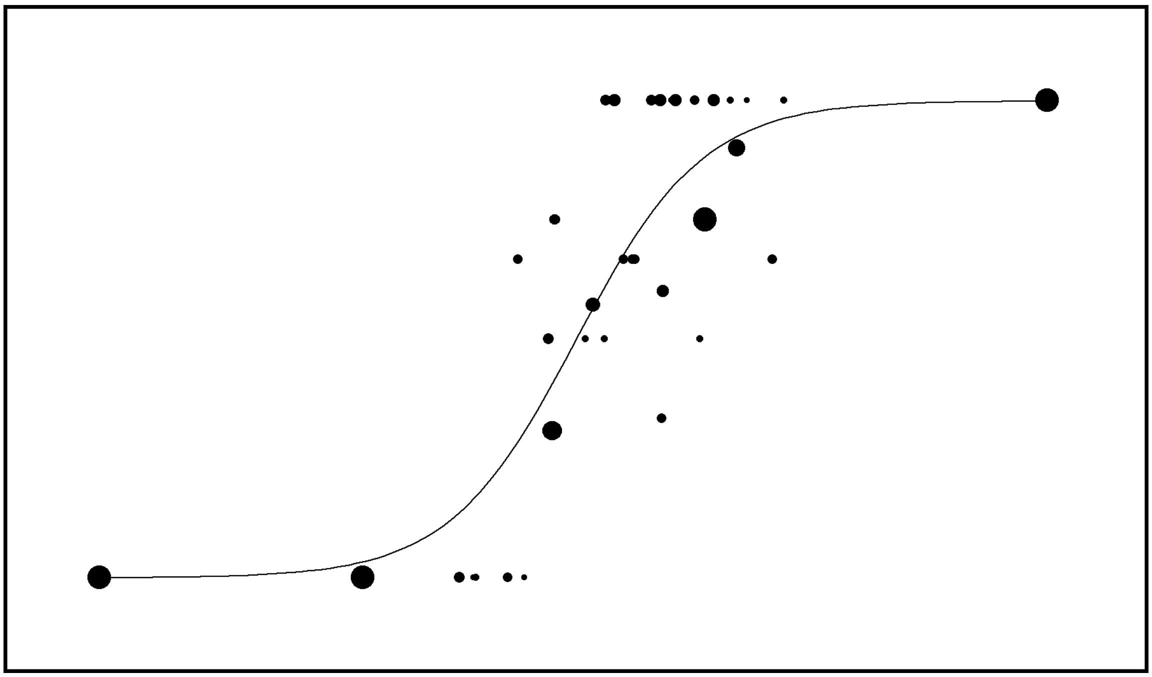


Figure C-4: Direct plot of data collected with adaptive routine.

We can solve the latter problem, that of having a different number of trials represented at each stimulus level, in two ways. If some kind of error bar is used in the plot, then this bar can be scaled with the inverse of the square of the number of trials; this is already implicit in the use of a standard deviation measure. The other method is to explicitly mark the data point with the number of trials. The size of a graphical symbol is easily scaled with the number of trials. For example, the area of a circular dot can represent the number of trials, such that the radius of a circular dot should be proportional to the square root of the number of trials in order that the ink will not misrepresent the data.

The first problem, that of the many stimulus levels used, is thornier. One way to plot the points would be to draw the best fitting curve first and then draw lines up or down from the curve depending on whether the trials at that point give a probability higher or lower than the curve at that level. The length of the line can be proportional to the standard deviation perhaps. There are two problems with this approach. The absolute scaling of the lengths of all the lines is a free parameter with great influence on the perception of whether the curve is a “good” fit or not. The scheme is also highly parametric, depending very much on the parameters of the fitted curve; it is unclear if a very bad fit would actually look particularly bad since the lines group around the curve.

A better way to deal with the multitude of levels is to group them into bins and thus form new pseudo-levels with many more trials. This procedure tends to smooth out the imaginary curve formed by the data-points. It also reduces the number of data-points plotted to the number of bins used. When the data is plotted in this way, it immediately clear whether the fit is reasonable or not, or phrased differently, whether the data set is reasonable or not. It only remains to establish how the number and size of the bins should be selected. A simple and consistent way to do this is to use *the best fitting function*. The width of bins can be specified in terms of the function parameters, for example some multiple of s when a logistic function is used, or s when a cumulative Gaussian is used. The position of the bins can be established by centering one bin at the middle of the fitted function, the other bins are then placed as well. Even though the data has been put into bins, the bins may still contain different numbers of trials. Therefore it is still a good idea to represent the number of trials by varying the size of the symbol used or in some other fashion. An example of a plot of the same data as shown in Figure C-5.

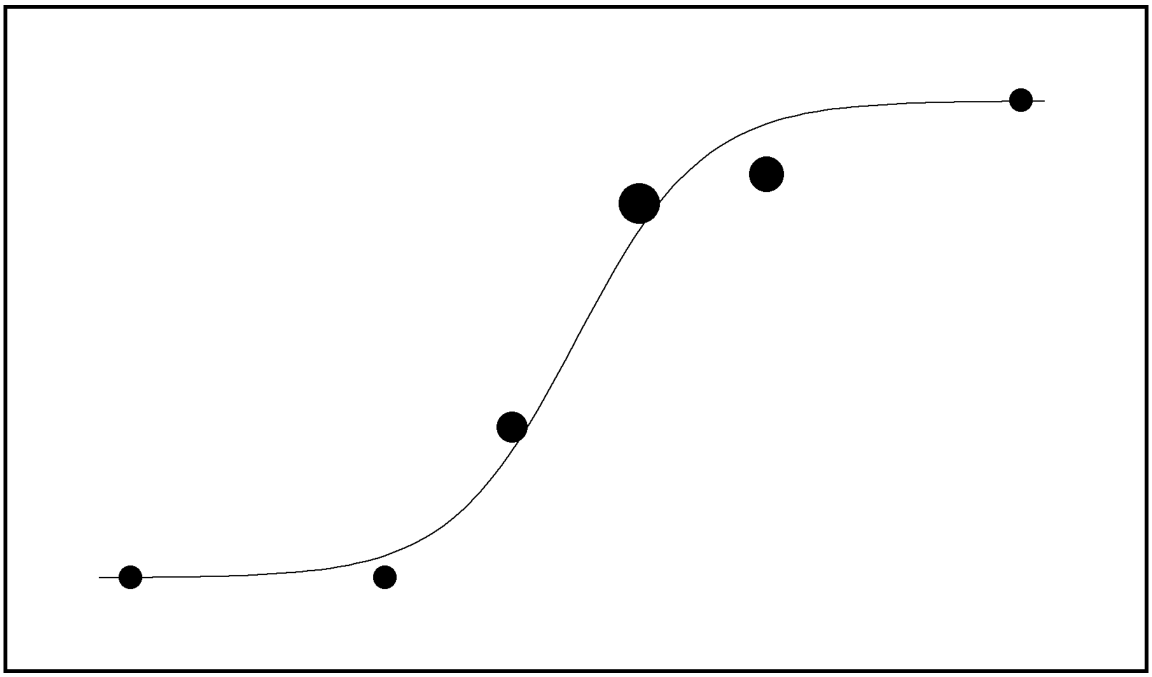


Figure C-5: A new plot of the same data in Figure C-4 using the binning method presented in the text.

This method is actually slightly parametric in that the widths and locations of bins are chosen based upon the paremeters of the best fitting function. Note that the position and width of the bins can be chosen in other ways, if there are good reasons for doing so. The main advantage of using the best fitting function is the ease with which different fits can be compared.

C.5 Bias Masquerading as Variability

Signal Detection Theory (SDT) postulates that there is an internal source of noise which is the cause of the variability observed in experiments. This source of noise is often assumed to be Gaussian, which is the reason why the psychometric function has the shape of a cumulative Gaussian function. These are reasonable assumptions and SDT works quite well. However, as this work has been concerned with finding the bias under various conditions The effects of bias have been considered carefully.

In an experiment there are some parameters which are varied and presented as different conditions; a good example of this type of parameter would be sizes of stimulus elements. Other parameters are treated as irrelevant and are randomly varied from trial to trial; an example of this would be the location of the entire stimulus on a display screen. These latter parameters can introduce variability and reduce the measured discriminability of the observer.

If the observers criterion changes position (bias) with changes in one of these presumed to be irrelevant parameters, then the responses given by subjects will be influenced by this parameter. This influence will not be consistent, since the parameter is varied randomly, and sometimes responses will contradict each other as the parameter is varied. This is how variability is increased. Simple calculations have shown that the influence of the randomly varied parameters will be reduced if a unifom distribution (or even better, a Gaussian distribution) is used. A narrower distribution will have a smaller effect. Also, if the shift in bias is on the order of the spread (s, inverse of slope) of the psychometric function, then its effect is negligible.

A more insidious version of the phenomenon occurs when an experimenter combines data points which are assumed to be symmetric about some midpoint when in fact the data points are not symmetric. In other words, if the experimenter assumes the location of the 50% point (in a Yes-No task) and then combines the data on both sides of this point, then the spread of the psychometric function will be overestimated. The mathematics of this case is almost identical to the case previously mentioned. Therefore, the effect is quite small if the misplacement is within one spread of the true psychophysical function. If the bias is large compared to the spread, then the estimated spread is proportional to the bias and not the actual spread.

The phenomenon of bias masquerading as variability can also be a problem in the 2AFC paradigm. Since two stimuli (signal and noise) are presented and the observer must determine which is which, the order of the stimuli must be randomized. However, if the observer has a preference for one stimuli or the other such that the response ratio is not 50%–50% but perhaps 60%–40%, then there is definitely "bias". This bias is not the same as the bias towards false-alarms or misses which is provided by SDT. The response preference mentioned should be accounted for before the SDT analysis in order to arrive at the correct d' and β. Unfortunately it is unclear exactly how this should be done.

The preference for one response over another is especially prevalent when the stimuli are presented serially. This was discovered in the early days of psychophysics using the method of constant stimuli with a standard, and is called the "time-error". One of the successes of *Gestalt* psycholgy was an explanation of the time-error with predictions which held up under experiments (Link, 1992).

It is likely that more can be gained from a more detailed study these issues.

Bibliography

Ahissar, M. & Hochstein, S. (2004) The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Sciences, 8*(10), 457-464.

Andrews, D. P., Butcher, A. K., & Buckley, B. R. (1973) Acuities for spatial arrangement of line figures: Human and ideal observers compared. *Vision Research, 13,* 599-620.

Aubert, G. & Kornprobst, P. (2006) **Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations** (2nd ed. Vol. 147). New York: Springer.

Badcock, D. & Westheimer, G. (1985a) Spatial location and hyperacuity: Flank position within the centre and surround zones. *Spatial Vision, 1*(1), 3-11.

Badcock, D. & Westheimer, G. (1985b) Spatial location and hyperacuity: The centre/surround localization contribution function has two substrates. *Vision Research, 25,* 1259-1267.

Berliner, A. & Berliner, S. (1948) The distortion of straight and curved lines in geometric fields. *American Journal of Psychology, 61,* 153-166.

Berry, R. N. (1948) Quantitative relations among vernier, real depth, and stereoscopic depth acuities. *Journal of Experimental Psychology, 38,* 708-721.

Best, F. (1900) Über die Grenze der Erkennbarkeit von Lagenunterschieden. *Albrecht von Graefes Archiv für Ophthalmologie, 51,* 453-460.

Bhatt, R., Carpenter, G. A., & Grossberg, S. (2007) Texture segregation by visual cortex: perceptual grouping, attention, and learning. *Vision Research, 47*(25), 3173-3211.

Bradley, A. & Freeman, R. D. (1985) Is reduced vernier acuity in amblyopia due to position contrast, or fixation deficits? *Vision Research, 25,* 55-66.

Bradley, A. & Skottun, B. C. (1987) Effects of contrast and spatial frequency on vernier acuity. *Vision Research, 27*(10), 1817-1824.

Burbeck, C. A. (1987) Position and spatial frequency in large-scale localization judgments. *Vision Research, 27*(3), 417-427.

Campbell, F. W. & Gubisch, R. W. (1966) Optical quality of the human eye. *The Journal of Physiology, 186*(3), 558.

Carandini, M. & Heeger, D. J. (1994) Summation and division by neurons in primate visual cortex. *Science, 264*(5163), 1333-1336.

Carlson, C. R. & Klopfenstein, R. W. (1985) Spatial-frequency model for hyperacuity. *Journal of the Optical Society of America A, 2*(10), 1747-1751.

Carpenter, G. A. (1989) Neural network models for pattern recognition and associative memory. *Neural Networks, 2*(4), 243-257.

Coombes, S. (2005) Waves, bumps, and patterns in neural field theories. *Biological Cybernetics, 93*(2), 91-108.

Daugman, J. G. (1980) Two-dimensional spectral analysis of cortical receptive field profiles. *Vision Research, 20,* 847-856.

Daugman, J. G. (1985) Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *Journal of the Optical Society of America A, 2*(7), 1160-1169.

Daugman, J. G. (2007) New methods in iris recognition. *IEEE Transactions on Systems, Man, and Cybernetics, Part B, 37*(5), 1167-1175.

Day, R. H. (1962) Excitatory and inhibitory processes as the basis of contour shift and negative after-effect. *Psychologia, 5,* 185-193.

De Valois, R. L. & De Valois, K. K. (1991) Vernier acuity with stationary moving Gabors. *Vision Research, 31*(9), 1619-1626.

Fahle, M., Edelman, S., & Poggio, T. (1995) Fast perceptual learning in hyperacuity. *Vision Research, 35*(21), 3003-3013.

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

Finkel, L. H. & Edelman, G. M. (1989) Integration of distributed cortical systems by reentry: a computer simulation of interactive functionally segregated visual areas. *Journal of Neuroscience, 9,* 3188-3208.

Finney, D. J. (1952) **Probit analysis: a statistical treatment ofthe sigmoid response curve.** (2nd ed.). Cambridge: Cambridge University Press.

Finney, D. J. (1971) **Probit Analysis.** Cambridge: Cambridge University Press.

Foster, D. H. & Bischof, W. F. (1991) Thresholds From Psychometric Functions: Superiority of Bootstrap to Incremental and Probit Variance Estimators. *Psychological Bulletin, 109*(1), 152-159.

Francis, G. (1993) *Cortical models of visual perception: Dynamics of form and motion segmentation*. Boston University, Boston, MA.

French, J. W. (1920) The unaided eye; part III. *Transactions of the Optical Society, 21*, 127-156.

Gabor, D. (1946) Theory of communication. *Journal of the Institute of Electrical Engineers, 93*, 429-457.

Geisler, W. S. (1984) Physical limits of acuity and hyperacuity. *Journal of the Optical Society of America A, 1*(7), 775-782.

Gove, A., Grossberg, S., & Mingolla, E. (1995) Brightness perception, illusory contours, and corticogeniculate feedback. *Visual Neuroscience, 12*(06), 1027-1052.

Grady, L. J. & Polimeni, J. R. (2010) **Discrete Calculus: Applied Analysis on Graphs for Computational Science.** New York: Springer.

Graham, N., Beck, J., & Sutter, A. (1992) Nonlinear processes in spatial-frequency channel models of perceived texture segregation: effects of sign and amount of contrast. *Vision Research, 32*(4), 719-743.

Grossberg, S. (2003) Laminar cortical dynamics of visual form perception. *Neural Networks, 16*(5), 925-932.

Grossberg, S. (2007) Towards a unified theory of neocortex: laminar cortical circuits for vision and cognition. *Progress in Brain Research, 165*, 79-104.

Grossberg, S. & Yazdanbakhsh, A. (2005) Laminar cortical dynamics of 3D surface perception: Stratification, transparency, and neon color spreading. *Vision Research, 45*(13), 1725-1743.

Grossberg, S. (1973) Contour enhancement, short-term memory, and constancies in reverberating neural networks. *Studies in Applied Mathematics, 52*, 217-257.

Grossberg, S. (1983) The quantized geometry of visual space: The coherent computation of depth, form and lightness. *The Behavioral and Brain Sciences, 6*, 625-692.

Grossberg, S. (1987) Cortical dynamics of three-dimensional form, color, and brightness perception: I. Monocular theory. *Perception & Psychophysics, 41*(2), 87-116.

Grossberg, S. & Mingolla, E. (1985a) Neural Dynamics of Form Perception: Boundary Completion, Illusory Figures, and Neon Color Spreading. *Psychological Review, 92*(2), 173-211.

Grossberg, S. & Mingolla, E. (1985b) Neural dynamics of perceptual grouping: Textures, boundaries, and emergent segmentations. *Perception & Psychophysics, 38*(2), 141-171.

Grossberg, S. & Mingolla, E. (1987) Neural dynamics of surface perception: Boundary webs, illuminants, and shape-from-shading. *Computer Vision, Graphics, and Image Processing, 37*, 116-165.

Grossberg, S., Mingolla, E., & Ross, W. D. (1997) Visual brain and visual perception: How does the cortex do perceptual grouping? *Trends in Neuroscience, 20*(3), 106-111.

Grossberg, S. & Rudd, M. E. (1992) Cortical dynamics of visual motion perception: Short-range and long-range apparent motion. *Psychological Review, 99*, 78-121.

Harris, J. P. & Fahle, M. (1996) Differences between fovea and periphery in the detection and discrimination of spatial offsets. *Vision Research, 36*(21), 3469-3477.

Heeger, D. J. (1992) Normalization of cell responses in cat striate cortex. *Visual Neuroscience, 9*(2), 181-197.

Hering, E. (1899) Über die Grenzen der Sehschärfe. *Berichte über die Verhandlungen der Königlich-Sächsischen Gesellschaft der Wissenschaften zu Leipzig, Mathematisch-Physische Klasse; Naturwissenschaftlicher Teil,* 16-24.

Hirsch, J. & Mjolsness, E. (1992) A center-of-mass computation describes the precision of random dot displacement discrimination. *Vision Research, 32*(2), 335-346.

Ijspeert, J. K., Van Den Berg, T., & Spekreijse, H. (1993) An improved mathematical description of the foveal visual point spread function with parameters for age, pupil size and pigmentation. *Vision Research, 33*(1), 15-20.

Jastrow, J. (1891) A study of Zoellner's figures and other related illusions. *American Journal of Psychology, 4*, 381-398.

Kellman, P. J. & Shipley, T. F. (1991) A theory of visual interpolation in object perception. *Cognitive Psychology, 23*(2), 141-221.

Kelly, F. & Grossberg, S. (2000) Neural dynamics of 3-D surface perception: Figure-ground separation and lightness perception. *Perception and Psychophysics, 62*(8), 1596-1618.

Klein, S. A. & Levi, D. M. (1985) Hyperacuity thresholds of 1 sec: Theoretical predictions and empirical validation. *Journal of the Optical Society of America A, 2*(7), 1170-1190.

Koenderink, J. J. (1984) The concept of local sign. In *Van Doorn, A. J., W. A. Van de Grind, & Koenderink, J. J.* (Eds.), *Limits in Perception*. Utrecht: VNU Science Press.

Kooi, F. L., De Valois, R. L., & Switkes, E. (1991) Spatial localization across channels. *Vision Research, 31*(9), 1627-1631.

Lee, T. S. (1996) Image representation using 2D Gabor wavelets. *Pattern Analysis and Machine Intelligence, IEEE Transactions on, 18*(10), 959-971.

Lesher, G. W. (1995) Illusory contours: Toward a neurally based perceptual theory. *Psychonomic Bulletin and Review, 2*, 279-321.

Lesher, G. W. & Mingolla, E. (1993) The role of edges and line-ends in illusory contour formation. *Vision Research, 33*(16), 2253-2270.

Levi, D. M., Klein, S. A., & Aitsebaomo, A. P. (1985) Vernier acuity, crowding and cortical magnification. *Vision Research, 25*, 963-977.

Levi, D. M. & Waugh, S. J. (1996) Position acuity with opposite-contrast polarity features: Evidence for a nonlinear collector mechanism for position acuity? *Vision Research, 36*(4), 573-588.

Link, S. (1992) **The Wave Theory of Difference and Similarity.** Hillsdale, NJ: Lawrence Erlbaum Associates.

Liu, J., Lu, Z. L., & Dosher, B. A. (2010) Augmented Hebbian reweighting: Interactions between feedback and training accuracy in perceptual learning. *Journal of Vision*.

Luce, R. D. (1959) **Individual Choice Behavior: A theoretical analysis.** New York: Wiley.

Luce, R. D., Bush, R. R., & Galanter, E. (Eds.). (1963a) **Readings in Mathematical Psychology, Vol. 1.** New York: Wiley.

Luce, R. D., Bush, R. R., & Galanter, E. (Eds.). (1963b) **Handbook of Mathematical Psychology, Vols. 1 & 2.** New York: Wiley.

Ludvigh, E. (1953) Direction sense of the eye. *American Journal of Ophthalmology, 36*, 139-142.

Macmillan, N. A. & Creelman, C. D. (1990) **Detection Theory: A user's guide.** New York: Cambridge University Press.

Marčelja, S. (1980) Mathematical description of the responses of simple cortical cells. *Journal of the Optical Society of America, 70*(11), 1297-1300.

Marr, D. (1982) **Vision.** San Francisco: W. H. Freeman.

McKee, S. P. & Westheimer, G. (1978) Improvement in vernier acuity with practice. *Perceptual Psychophysics, 24*(3), 258-262.

Morgan, M. J., Ward, R. M., & Hole, G. J. (1990) Evidence for positional coding in hyperacuity. *Journal of the Optical Society of America A, 7*(2), 297-304.

Nakayama, K., Shimojo, S., & Silverman, G. H. (1989) Stereoscopic depth: its relation to image segmentation, grouping, and the recognition of occluded objects. *Perception, 18*(1), 55-68.

Nothdurft, H.-C. (1992) Feature analysis and the role of similarity in preattentive vision. *Perception & Psychophysics, 52,* 355-375.

Parent, P. & Zucker, S. W. (1989) Trace inference, curvature consistency, and curve detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on, 11*(8), 823-839.

Pentland, A. (1980) Maximum likelihood estimation: The best PEST. *Perception & Psychophysics, 28,* 377-379.

Peterhans, E., von der Heydt, R., & Baumgartner, G. (1986) Neuronal responses to illusory contour stimuli reveal stages of visual cortical processing. In J. D. Pettigrew, K. J. Sanderson, & W. R. Levick (Eds.), *Visual Neuroscience*. Cambridge: Cambridge University Press.

Petrov, A. A., Dosher, B. A., & Lu, Z.-L. (2005) The dynamics of perceptual learning: An incremental reweighting model. *Psychological Review, 112,* 715-743.

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.

Poggio, G. F., Doty, J., R. W., & Talbot, W. H. (1977) Foveal striate cortex of behaving monkey: Single-neuron responses to square-wave gratings during fixation of gaze. *Journal of Neurophysiology, 40,* 1369-1391.

Poggio, T., Fahle, M., & Edelman, S. (1992) Fast perceptual learning in visual hyperacuity. *Science, 256*(5059), 1018-1021.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (1986) **Numerical Recipes.** Cambridge: Cambridge University Press.

Resnikoff, H. L. (1989) **The illusion of reality.** New York: Springer-Verlag.

Rivest, J. & Cavanagh, P. (1996) Localizing contours defined by more than one attribute. *Vision Research, 36*(1), 53-66.

Ruda, H. (1998) The warped geometry of visual space near a line assessed using a hyperacuity displacement task. *Spatial Vision, 11*(4), 401-419.

Saarinen, J. & Levi, D. M. (1995) Perceptual learning in vernier acuity: what is learned? *Vision Research, 35*(4), 519-527.

Schumann, F. (1987) Contributions to the Analysis of Visual Perception – First Paper: Some Observations on the Combination of Visual Impressions into Units. (A. Hogg, Trans.). In S. Petry & Meyer, G, E (Eds.), *The Perception of Illusory Contours*. (pp. 21-34). New York: Springer-Verlag.

Schumann, F. (1900) Beiträge zur Analyse der Gesichtswahrnehmungen. Erste Abhandlung. Einige Beobachtungen über die Zusammenfassung von Gesichtseindrücken zu Einheiten. *Zeitschrift für Psychologie und Physiologie der Sinnesorgane, 23*, 1-32.

Smith, P. L. & Vickers, D. (1988) The accumulator model of two-choice discrimination. *Journal of Mathematical Psychology, 32*(2), 135-168.

Soriano, M., Spillmann, L., & Bach, M. (1996) The abutting grating illusion. *Vision Research, 36*(1), 109-116.

Sotiropoulos, G., Seitz, A. R., & Series, P. (2011) Perceptual learning in visual hyperacuity: A reweighting model. *Vision Research, 51*(6), 585-599.

Spratling, M. W. (2010) Predictive coding as a model of response properties in cortical area V1. *Journal of Neuroscience, 30*(9), 3531-3543.

Sutter, A., Beck, J., & Graham, N. (1989) Contrast and spatial variables in texture segregation: testing a simple spatial-frequency channels model. *Perception & Psychophysics, 46*(4), 312-332.

Taylor, M. M. (1971) On the efficiency of psychophysical measurement. *Journal of the Acoustical Society of America, 49*, 505-508.

Toet, A. & Koenderink, J. J. (1988) Differential spatial displacement discrimination thresholds for Gabor patches. *Vision Research, 28*(1), 133-143.

Tyler, C. W. (1973) Periodic vernier acuity. *Journal of Physiology, 228*, 637-647.

van den Berg, T. J., Franssen, L., & Coppens, J. E. (2009) Straylight in the human eye: testing objectivity and optical character of the psychophysical measurement. *Ophthalmic and Physiological Optics, 29*(3), 345-350.

Vickers, D. (1970) Evidence for an accumulator model of psychophysical discrimination. *Ergonomics, 13*(1), 37-58.

Vickers, D. (1979) **Decision processes in visual perception.** Academic Press New York:.

Vickers, D. & Lee, M. D. (1998) Dynamic models of simple judgments: I. Properties of a self-regulating accumulator module. *Nonlinear Dynamics, Psychology, and Life Sciences, 2*(3), 169-194.

Vickers, D. & Lee, M. D. (2000) Dynamic models of simple judgments: II. Properties of a self-organizing PAGAN (parallel, adaptive, generalized accumulator network) model for multi-choice tasks. *Nonlinear Dynamics, Psychology, and Life Sciences, 4*(1), 1-31.

Vickers, D. & Smith, P. (1985) Accumulator and random-walk models of psychophysical discrimination: a counter-evaluation. *Perception, 14*(4), 471-497.

Volkmann, A. W. (1863) **Physiologishe Untersuchungen im Gebiete der Optik.** Leipzig: Breitkopf & Härtel.

Watson, A. B. & Pelli, D. G. (1983) QUEST: A Bayesian adaptive psychometric method. *Perception & Psychophysics, 33*(2), 113-120.

Watt, R. J. & Andrews, D. P. (1981) APE: Adaptive probit estimation of psychometric functions. *Current Psychological Review, 1,* 205-214.

Watt, R. J. & Campbell, F. W. (1985) Vernier acuity: Interactions between length effects and gaps when orientation cues are eliminated. *Spatial Vision, 1*(1), 31-38.

Waugh, S. J., Levi, D. M., & Carney, T. (1993) Orientation, masking, and vernier acuity for line targets. *Vision Research, 33*(12), 1619-1638.

Waugh, S. J. & Levi, D. M. (1995) Spatial alignment across gaps: contributions of orientation and spatial scale. *Journal of the Optical Society of America A, 12*(10), 2305-2317.

Weiss, Y., Edelman, S., & Fahle, M. (1993) Models of perceptual learning in vernier hyperacuity. *Neural Computation, 5*(5), 695-718.

Westheimer, G. (1975) Editorial: visual acuity and hyperacuity. *Investigative Ophthalmology, 14,* 570-572.

Westheimer, G. (1978) Spatial phase sensitivity for sinusoidal grating targets. *Vision Research, 18,* 1073-1074.

Westheimer, G. (1979) The spatial sense of the eye. *Investigative Ophthalmology & Visual Science, 18*(9), 893-912.

Westheimer, G. (1981) Visual hyperacuity. *Progress in Sensory Physiology, 1,* 1-30.

Westheimer, G. (1984) Line-separation discrimination curve in the human fovea: Smooth or segmented? *Journal of the Optical Society of America A, 1*(6), 683-684.

Westheimer, G. (2005) Anisotropies in peripheral vernier acuity. *Spatial Vision, 18*(2), 159-167.

Westheimer, G. & Hauske, G. (1975) Temporal and spatial interference with vernier acuity. *Vision Research, 15*, 1137-1141.

Westheimer, G. & McKee, S. P. (1977a) Integration regions for visual hyperacuity. *Vision Research, 17*, 89-93.

Westheimer, G. & McKee, S. P. (1977b) Spatial configurations for visual hyperacuity. *Vision Research, 17*, 941-947.

Weymouth, F. W., Andersen, E. E., & Averill, H. L. (1923) Retinal mean local sign: A new view of the relation of the retinal mosaic to visual perception. *American Journal of Physiology, 63*, 410-411.

Whitaker, D. & MacVeigh, D. (1991) Spatial frequency and separation in vernier acuity. *Vision Research, 31*(7/8), 1205-1212.

Williams, R. A. & Essock, E. A. (1986) Areas of spatial interaction for a hyperacuity stimulus. *Vision Research, 26*(2), 349-360.

Wilson, H. R. (1986) Responses of spatial mechanisms can explain hyperacuity. *Vision Research, 26*(3), 453-469.

Wilson, H. R. (1991) Model of peripheral and amblyopic hyperacuity. *Vision Research, 31*(6), 967-982.

Wilson, H. R. & Gelb, D. J. (1984) Modified line-element theory for spatial-frequency and width discrimination. *Journal of the Optical Society of America A, 1*(1), 124-131.

Wilson, H. R., McFarlane, D. K., & Phillips, G. C. (1983) Spatial frequency tuning of orientation selective units estimated by oblique masking. *Vision Research, 23*(9), 873-882.

Wülfing, E. A. (1892) Über den kleinsten Gesichtswinkel. *Zeitschrift Für Biologie, 29*, 199-202.

Yellott Jr, J. I. (1983) Spectral consequences of photoreceptor sampling in the rhesus retina. *Science, 221*(4608), 382.

Yu, C. & Essock, E. A. (1996) Psychophysical end-stopping associated with line targets. *Vision Research, 36*(18), 2883-2896.

Zhou, H., Friedman, H. S., & Von Der Heydt, R. (2000) Coding of border ownership in monkey visual cortex. *The Journal of Neuroscience, 20*(17), 6594-6611.

Curriculum Vitae

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]