SPIE Proceedings 3016 "Human Vision and Electronic Imaging II" eds. Rogowitz & Pappas, (1997), p. 13 - 24.
Invited paper

Seven models of masking

Stanley A. Klein[a], Thom Carney[a,b], Lauren Barghout-Stein[a,c] and Christopher W. Tyler[c]
[a] School of Optometry, UC Berkeley, Berkeley, CA 94720
[b] Neurometrics Institute, Berkeley, CA 94619
[c] Smith Kettlewell Eye Research Institute, San Francisco, CA 94115

ABSTRACT

Seven types of masking are discussed: multi-component contrast gain control, one-component transducer saturation, two-component phase inhibition, multiplicative noise, high spatial frequency phase locked interference, stimulus uncertainty, and noise intrusion. In the present vision research community, multi-component contrast gain is gaining in popularity while the one- and two-component masking models are losing adherents. In this paper we take the presently unpopular stance and argue against multi-component gain control models. We have a two-pronged approach. First, we discuss examples where high contrast maskers that overlap the test stimulus in both position and spatial frequency nevertheless produce little masking. Second, we show that alternatives to gain control are still viable, as long as uncertainty and noise intrusion effects are included. Finally, a classification is offered for different types of uncertainty effects that can produce large masking behavior.

Keywords: human vision, model, masking, gain control, uncertainty, noise

INTRODUCTION

Models of masking are important for many reasons. On a fundamental level, when one understands masking one will have gained a deep understanding of pattern discrimination and visual processing in general. On a practical level, a successful model of visual masking will directly improve image compression by being able to save image bits where they are masked and thus not needed.

In a general masking situation one is asked to detect a test pattern in the presence of a suprathreshold masking pattern. In its application to medical imaging, the test pattern could be a tumor and the masker would be the noisy image of normal tissue. In its application to image compression our goal is to develop a fidelity metric to predict when the compressed and decompressed image is discriminable from the original image. The original image would be the mask and the difference image would be the test. The fidelity metric that is being developed is a vision model that measures the visibility of this test image in the presence of the masker.

Models of masking as developed by the vision community may at first be thought to have limited applicability for image processing applications. The vision community typically uses simple stimuli that are easily memorized and usually generate limited amounts of masking. With complex stimuli such as found in real scenes, on the other hand, the test pattern (the distortion error produced by compression) and the pedestal (the full scene) may not be familiar and the amount of masking can be much larger. The basic vision community is increasingly becoming interested in the more difficult aspects of real world scenes, partly because our visual system has evolved to process complex visual stimuli, so we may be neglecting important properties of visual processing if we only examine simple stimuli. A second reason for basic researchers to use complex stimuli is that improved vision models are needed in the applied areas, and we can contribute to that effort. As the stimuli get more complex, uncertainty effects become more important.

This article examines seven types of masking:

<u>Type 1. Pooled contrast gain control</u>. Albrecht and Geisler[1] and Heeger[2] have developed an elegant contrast gain control (divisive inhibition) formalism that seems to be in agreement with physiology. In recent years this approach has become popular for psychophysical modeling. Its increasing popularity encourages us to play the devil's advocate and question the generality of pooled contrast gain. We will discuss evidence for why we believe that pooled contrast gain control plays only a small role in psychophysics and why we believe that alternative models may be able to do a better job than pooled divisive inhibition. The problem with past efforts to discriminate between masking models is that one needs to explicitly lay out the best versions of competing models for comparison.

<u>Type 2. Transducer saturation or one-component pooling</u>. This model advocated by Legge & Foley[3] and Wilson[4] has fallen out of favor for a number of reasons, to be discussed. These arguments are not fully convincing to us and a modified transducer saturation model is proposed (Type 3) as a viable candidate for a broad range of masking experiments.

<u>Type 3. Phase inhibition or Pythagorean (two component) pooling</u>. Klein & Stromeyer[5] suggested that inhibition between different phases may be important. Klein & Levi[6] suggested that mechanisms with different phases are summed together

before they go though a saturating nonlinearity. This two-component model is intermediate between the many-component (Type 1) and the one-component (Type 2) pooling models listed above.

Type 4. Multiplicative noise. The idea here is that the mechanism's noise variance is proportional to the stimulus strength[7]. This type of masking makes predictions that are similar to those of the other masking types. We will suggest several ways in which this possibility can be discriminated from the others.

Type 5. High spatial frequency phase-locked patterns. We will describe several maskers that produce more masking than standard filter models predict. These patterns include masking by beats and by sampled stimuli. However, since these phenomena are not yet well understood we will not discuss them in detail.

Type 6. Stimulus uncertainty. This is the category in which the observer monitors the channels stimulated by the mask in addition to those stimulated by the test. There are many occasions when the observer is unable to restrict attention to the relevant stimulus. Effects concerned with the observer's ability to focus attention may well be the most important, especially for real world images. They are also the most ignored of the masking models.

Type 7. Intrusive noise. This differs from type 6 masking (uncertainty) in that the noise actually does intrude into the test mechanism (not just because of observer uncertainty) thereby producing a deficit in the ideal observer.

We now discuss the seven types of masking in more detail.

1. POOLED CONTRAST GAIN CONTROL.

Partly because of the models of Heeger[2] and Albrecht & Geisler[1] and the physiological experiments of Albrecht & Geisler[8], and Ohzawa, Freeman & Sclar[9] the topic of contrast gain control has become very popular in the past few years. Malik & Perona[10], found an important computational role for contrast gain control in their successful model of texture segmentation. Wilson & Humanski[11], Foley[12], Thomas & Olzak[13], and Wilson & Kim[14] have recently provided further evidence for a contrast gain mechanism. It has become the model of choice for explaining masking and one of the goals of this paper is to raise questions about it. We will also examine alternative models.

By contrast gain control, we mean a model wherein the mask affects the test mechanism's response through a divisive inhibition pool of activity:

$$R(S_m, S_t) = T(S_t+\alpha S_m) / (1 + \Sigma_i M_i(S_t + \beta_i S_m)) \qquad (1)$$

where S_m and S_t are the stimulus strengths of the mask and test and R is the mechanism response. T and M_i stand for test and mask. T represents the direct activity of the optimal mechanism for detecting the test pattern, and M_i is the contribution of the ith masking mechanism to the gain-control pool. The functions T and M_i are nonlinear function that account for threshold effects, but not saturation effects. The saturation is achieved by the denominator. The summation is over many mechanisms, some that respond to the test (one of the M_i terms will be proportional to the numerator) and many that only respond to the mask.

The extent of the summation in the denominator is what distinguishes masking types 1, 2 and 3. We are calling Type 1 masking "pooled masking" because it involves a broad pool. Type 2 masking, often referred to as transducer saturation, will also be called single-component masking since the summation is restricted to just the test mechanism. Type 3 masking is similar to Type 2 except that the pool extends over mechanisms with different polarities.

In the Type 1 pool, the mechanisms can be at the same location as the test, but outside the orientation and spatial frequency tuning of the mechanism in the numerator. In addition, there can be contributions to M_i from mechanisms at other locations. The latter case will be called lateral contrast gain control, and will be investigated in another paper in this volume[15]. The parameter α in Eq. 1, is introduced to represent the test mechanism's relative sensitivity to the mask and test stimulus. Similarly β_i specifies the relative sensitivity for the ith mechanism in the gain-control pool. The summation in the denominator of Eq. 1 is depicted as linear but it could be a complicated nonlinear function

The response function given by Eq. 1 is often called the transducer function. It is connected to the psychophysically measured, signal detection discriminability index, d', by:

$$d' = R(S_m, S_t) - R(S_m, 0). \qquad (2)$$

where we have assumed the noise at the output has unit variance to get d' units. It is common to define the threshold test strength to be that strength which gives d'=1. For the case of no masking ($S_m = 0$) the second term in Eq. 2 vanishes. Eq. 2 is based on the assumption that there is no signal uncertainty. That is, the observer is only attending to the mechanisms that contain the signal information. If uncertainty is present then d' would be attenuated beyond what is given by Eq. 2 as will be discussed in Type 6 masking.

We don't doubt that Type 1 contrast gain control exists, we just question whether it produces very strong masking. As mentioned in the Introduction, pooled gain-control models are gaining in popularity. In the remainder of this section we discuss several discrimination experiments with high contrast maskers that show type 1 pooled gain-control effects are weak. We first start with an overview of what we call the test-pedestal framework.

Klein and colleagues have been finding that a wide range of hyperacuity discrimination data can be well understood in terms of a test-pedestal framework. Most of this work was reviewed in detail in the SPIE Proceedings on "Computational Vision Based on Neurobiology"[16]. Several of the experiments are listed in the following table. In general the threshold vs. contrast (tvc) curves that are found using these easily learned and memorized stimuli have shallow log-log slopes (<0.3) indicating that masking is weak. Since these test-pedestal results were reviewed by Klein[16] only a brief summary will be presented here.

	Pedestal	Test	facilitation	masking	Comment
contrast	cos(fx)[7,17]	cos(fx)	strong	moderate	The standard dipper function
	cos(fx)[7]	cos(3fx)	strong	weak	Multiple edge blur discrimination
motion	cos(fx)cos(ωt)[18]	sin(fx)sin(ωt)	very strong	weak	Detection of motion with flickering mask
Vernier	L(x)[19,20]	D(x)E(y)	none	weak	Line vernier acuity
	E(x)[19]	L(x)E(y)	"	"	Edge vernier acuity
	cos(fx)[21]	sin(fx)E(y)	none	weak	Sinusoidal vernier acuity
bisection	L(x)[6,22]	D(x)	none	weak	Three- and five-line bisection
jitter	cos(fx)[23]	sin(fx)sin(ωt)	none	variable	Back and forth oscillations of a grating
flicker	cos(fx)[23]	cos(fx)cos(ωt)	weak	variable	Flickering contrast of a grating
resolution	E(x)[20]	D(x)	strong	weak	Single edge blur discrimination
(blur)	L(x)[20]	Q(x)	"	"	Single line blur discrimination

TABLE 1

The notation E(x), L(x), D(x), Q(x) are the multiple family of edge, line, dipole and quadrupole. Each is the derivative of the preceding one. We now consider several of these experiments in more detail.

1.1. Blur discrimination. Recently we have extended the test-pedestal approach to blur (resolution) discrimination using a high contrast edge (or line) as the pedestal and a dipole (or quadrupole) as the test[20]. When expressed in these multipole units, we find that the resolution thresholds are typically at or below the test detection threshold. This is lower (better) than the threshold for detecting that multipole in the vernier (hyperacuity) configuration. In the resolution case, even at very high pedestal contrasts the test threshold elevation was very small. Thus there was minimal masking. Our finding turns upside-down the usual belief that hyperacuity thresholds are lower than resolution thresholds. When thresholds are expressed in arc minutes, then hyperacuity indeed has lower thresholds. However, we believe that arc minutes are inappropriate units for comparing thresholds of dissimilar tasks. Better is to compare the thresholds for detecting the identical test pattern in the context of the different pedestals. One might worry about the legitimacy of comparing thresholds across different pedestals. We can offer three reasons why we feel it is legitimate: 1. Different pedestal strengths can be expressed in terms of pedestal detection threshold units. 2) Thresholds are fairly independent of pedestal contrast (especially for the resolution task). 3) The test-pedestal predictions are similar to what detailed filter models predict, since the optimal filter would be tuned for the test pattern while avoiding the mask. The important point in the present context is that there is minimal masking in a number of tasks such as blur discrimination even though the masker (pedestal) overlaps the test in both position and spatial frequency.

1.2. Bisection data imply weak gain control effects. Klein[16,24] pointed out that the low bisection thresholds for closely spaced stimuli are evidence against strong masking by a gain control mechanism. Here is how that argument goes. Consider the three-line bisection with line separations of 1.3 min[6]. The bisection threshold is about 1.2 sec (corresponding to a displacement of about 3/8 inch at a distance of a mile). A 5-line bisection task gives even smaller thresholds[25]. Carney & Klein[22] extended these measurements by obtaining bisection thresholds as a function of contrast. The tvc functions were quite shallow and the dipole detection threshold is an excellent predictor of the bisection threshold. Figure 13 of Klein & Levi[6] shows "viewprint" diagrams of the amount of stimulation of mechanisms with different peak spatial frequencies and at different center locations. The figure shows that 99% of the mechanisms are responding strongly with a differential response that precludes them from being responsible for the discrimination. Only a tiny fraction of the mechanisms are properly positioned in spatial frequency (zero-crossings matched to the stimulus separation) and location (between the lines) to carry out the discrimination. Klein[16] showed that the surrounding highly stimulated mechanisms reduced the gain of the detection mechanisms by at most a factor of about 3. This is not a large effect.

1.3. Opponent motion. Stromeyer, et al.[18] used a high contrast counterphase flickering grating as the masker and an identical grating either in-phase or in quadrature phase, as the test (see Table 1). For the in-phase case, there was the standard contrast masking. However, for the quadrature case there was strong facilitation rather than masking, even with high mask contrasts. One would think that in a contrast gain control model with divisive inhibition the strongly stimulated motion mechanisms would raise thresholds for all phases of the counterphase test. Our data showed that there were special opponent motion mechanisms that were not desensitized by the masker. This result is similar to the special purpose mechanisms in the blur and bisection tasks just discussed. So even though physiology reveals divisive inhibition in the general neural population, psychophysical experiments can give very different results, since the psychophysical judgment can be based on a small population of mechanisms that are isolated from the divisive pool.

1.4. Masking outside the receptive field. The hallmark of pooled contrast gain is that because of the denominator in Eq. 1, a masker can produce threshold elevation outside the normal receptive field. This topic has become quite confusing recently with Foley[12] showing very large masking effects (threshold elevations of up to 10-fold) for maskers at right angles to the test.

We suspect that strong long-range masking only occurs for synchronous, brief (33 msec) test and masker, such as used by Foley[12].The onset transient might cause a general, broad, disturbance to the visual system that raises thresholds for all stimuli. It might also be related to the Crawford masking luminance saturation immediately following a background change. This may be a retinal effect that has broad spatial frequency and orientation tuning. The divisive inhibition model of Wilson & Kim[14] takes time to develop (maybe 30 msec) and doesn't kick in until after Foley's stimulus would already have turned off. Masking studies with longer durations show much narrower orientation tuning bandwidths[26,27,28,29]. The short duration may also induce extra uncertainty (see Section 6.1.3). The paper by Foley is very well done, but it might not be revealing information about contrast gain control.

1.5. Connection to physiology. The contrast gain control model seems able to explain a wide assortment of physiological data on receptive field properties of neurons[1,2]. We must be cautious about a direct application of the contrast gain formalism to psychophysics. The gain control models of Albrecht & Geisler[1] and Heeger[2] do a decent job of accounting for physiological data and for psychophysical suprathreshold orientation and spatial frequency[13] discriminations. But that isn't the topic of the present paper. We are concerned with detecting test patterns in the presence of a background pattern. In physiological studies it might indeed be true that maskers outside the classical receptive field can produce threshold elevations. Thus the bandwidths from masking appears to be broader than the true mechanism bandwidth. However, in psychophysical testing, as we have pointed out, masking effects can be elusive. Under some circumstances there is distant masking[12], but often the masking can be avoided by using special purpose mechanisms that evade the gain control. It may be difficult to find these special purpose mechanisms in a physiological study since they may be sparse.

2. SINGLE COMPONENT TRANSDUCER FUNCTION WITH SATURATING NONLINEARITY.

Transducer saturation models can be thought of as a special case of contrast gain control in which the sum over M_i is replaced by the numerator function. The response of a mechanism to the test plus mask can thus be written as the transducer function, R(S):

$$R(S) = R(S_t + \alpha S_m) \tag{3}$$

This transducer function is distinctive because it is a function of a single variable, $S = S_t + \alpha S_m$. That is why we call it single-component masking. A sample function that shows transducer masking would be a response function of the form.

$$R(S) = S^2 / (1 + S^{1.5}) \tag{4}$$

This function has a quadratic behavior at low stimulus strengths and square root behavior at high contrast. A similar function with logarithmic behavior at high contrast was proposed by Klein & Levi[6]. A number of authors have pointed out problems with the single-component model, as we now discuss.

2.1. Phase independence. Beard, Klein & Carney[23] modeled their data using a transducer function of the form given by Eq. 3. They showed how the transducer function approach can be used to fit a broad range of masking data. Their experiment used a static pedestal mask and a counterphase grating test pattern, either in-phase for a flickering stimulus or in quadrature phase for a jittering stimulus. They found that to first order, the masking was phase-independent, but with a small tendency for the jitter (quadrature) stimulus to be a bit more visible (less masking). A single-component transducer function isn't able to explain the phase independence, so we believe it should be replaced with a two-component energy model (see Type 3 masking below).

2.2. Off-frequency looking. One would think that the single-component restriction specified by Eq. 3 as to how the transducer function depends on the mask should make it easy to disprove the transducer model since for a given test pattern one might expect the optimal mechanism to be fixed. A fixed test mechanism would lead to very strong constraints on the expected results. However, the notion of 'off-frequency looking' allows different test mechanisms to be optimal for different maskers, making the simple story a bit more complicated. Klein[16] provided a detailed analysis of 'off-frequency looking' in analyzing the interactions of a third-harmonic masker and a first harmonic test (the Stromeyer & Klein[7] data). A bit of history may be helpful in providing background. One of the earliest papers on the spatial frequency approach to vision was by Campbell & Robson[30] showing that a square wave could be distinguished from a sine wave when the third harmonic just reaches threshold. In the context of image fidelity, that task was a blur resolution task in that a sine wave is a blurred square wave. The Campbell & Robson result provided a method for predicting the resolution threshold. Stromeyer & Klein[7] extended the first plus third harmonic experiments to higher pedestal contrasts and showed that the blur discrimination occurs at about half the third harmonic threshold. A model was developed for predicting this facilitation, similar to the dipper function found when the pedestal had the same frequency as the test[7,17]. An important feature of the Stromeyer & Klein[7] model is the notion of 'off-frequency looking'. The idea is that a change in the masker contrast alters the peak spatial frequency of the optimal mechanism for the discrimination task (see Klein[24] for details). By shifting the optimal mechanism as contrast increases, the amount of masking is reduced. The 'off-frequency looking' mechanism adds needed flexibility to a restrictive single- (Type 1) or double- (Type 3) component masking model.

2.3. Masking outside the receptive field. Another seemingly easy way to discount the transducer model is to find cases in which the masker doesn't stimulate the classical receptive field of the test as measured by adaptation or facilitation (such as an oriented mask stimulus at right angles to the test stimulus). For the gain control case, α (in Eq. 3) can be zero and yet the mask could still raise thresholds even though it doesn't directly excite the mechanism. There are two problems with this objection to saturating transducers.

First, we are suspicious about the generality of data that show large masking by distant maskers. As discussed earlier (Section 1.4), Foley[12] finds threshold elevations of up to 10-fold by maskers at right angles to the test. Foley used brief exposures to avoid adaptation effects, but the result of his brief presentation may have been to introduce saturation effects of the sort found in luminance Crawford masking[31,32]). Ironically, Foley may have adopted conditions that reveal saturation behavior of early luminance processing (retinal) stages, instead of the orientation selective cortical behavior. In gain-control models the denominator in Eq. 1 typically takes a short time to develop. If the test is presented before the gain control has set in then the test pattern will be saturated. If the test pattern is presented a bit later, then the denominator will grow in magnitude and desensitize the response to all stimuli, bringing the system out of saturation. Thus, Foley may have been measuring transducer saturation at a different processing stage than what he had intended. A stage of retinal compression followed by cortical facilitation and tuning may be a good model for fitting Foley's data.

Second, masking by a distant masker could be incorporated into a saturating transducer framework if there were a masker threshold. Instead of the mask dependence being $S = S_t + \alpha S_m$ as in Eq. 1, the response would be a function of S, given by:

$$\begin{aligned} S &= S_t && \text{for } S_m < th_m \\ &= S_t + \alpha (S_m - th_m) && \text{for } S_m \geq th_m \end{aligned} \tag{5}$$

where the modification from Eq. 1 sets the contribution of the mask to zero for mask strengths less than the mask threshold, th_m. This minor modification to the saturating transducer model allows it to be a viable contender for masking models, in that the masker wouldn't have any effect when using near threshold stimuli. It would only become important at high contrasts.

2.4. Predicted linearization. In order to calculate d' in a masking experiment, the differential response of a mechanism (or assembly of mechanisms) to the test pattern is needed. It is given by:

$$d' = T(S_t) = R(\alpha S_m + S_t) - R(\alpha S_m) \tag{6}$$

When the test strength is small relative to the mask the response is approximately:

$$d' \approx R'(\alpha S_m)\, S_t, \tag{7}$$

where R' is the derivative of R. If R' is a decreasing function of S (a saturating nonlinearity), then the visibility of the test is a decreasing function of S_m and we get masking[3].

Eq. 7 provides a possible test of divisive inhibition vs. transducer masking. Eq. 7 implies a linear dependence of d' on test strength, S_t. On the other hand, the divisive inhibition prediction (Eq. 3), would have d' maintaining its power function dependence on test strength. Kersten[33] and Legge, Kersten & Burgess[34] made use of this logic to provide evidence in favor of transducer masking as opposed to uncertainty or gain control masking. Legge et al.[34] measured detection and contrast discrimination of a Gabor function on a noise background. When the Gabor pedestal and the noise background were zero (standard detection), the transducer exponent was 2.06 and 2.13 for the two observers (i.e. $d' \approx c^{2.1}$). Small amounts of either the Gabor pedestal or the noise background resulted in a lowering of the exponent to between 0.9 and 1.5.

The results of Kersten[33] and Legge, et al.[34] seem to provide strong evidence in favor of a transducer type masking. Since these results are in agreement with our own notions, we feel obliged to cast some doubt on their generality. We would like to consider the possibility that their results are a consequence of stimulus uncertainty in the unmasked condition. When the mask wasn't present, there seemed to be a large amount of temporal uncertainty in the stimulus. In their experiments the mask and test overlapped in time (160 msec duration), so the mask reduced temporal uncertainty for the test presentation. If the mask had been on constantly then we believe the exponent would have stayed at around 2 because of temporal uncertainty even though the threshold would be elevated due to transducer saturation. Kersten[33] had one condition that tested this possibility, and the transducer exponent did increase from 1.0 to 1.4. This increase is in the right direction, but is too small, for the contrast-gain model. More data of this sort is needed in order to develop greater confidence that noise can linearize the transducer function.

In the companion paper by Barghout, Tyler & Klein[15] we examine masking by a surrounding annulus. The masking that is found cannot be explained by a fixed transducer function since we found that the transducer exponent stays at approximately 2 both with and without the masker.

3. PHASE INHIBITION AND PYTHAGOREAN PHASE POOLING.

We have carried out a number of studies that indicate phase pooling early in visual processing. Beard, Klein & Carney[23] compared flicker and jitter (see stimulus description in Table 1) where the test was a counterphase grating and the masker was a stationary sinusoid of the same spatial frequency as the test. The masker had either the same spatial phase as the test (flicker) or it was in quadrature phase (jitter). Our results were that, to first order, the flicker and jitter thresholds were about equal. Earlier evidence for phase independence was also found in comparing sinusoidal vernier acuity and contrast discrimination[21]. We interpret these data as evidence against a single-component gain control (Type 2 masking) and in favor of an energy model in which even and odd symmetric mechanisms are combined by a Pythagorean sum before further processing[6]. Beard et al.[23] used a response function proposed by Klein & Levi[6]:

$$R(S) = A \ln(1 + nwE^n) \qquad (8)$$

where n is the power function exponent (taken to equal 2) governing the low contrast behavior, w is the high contrast Weber fraction and E is the stimulus energy. Eq. 8 is similar to the combination of test and mask strengths suggested in Eq. 3, except it is based on E, the Pythagorean sum of even and odd symmetric mechanisms. This transducer function has the useful (and surprising) property that in using Eqs. 2,3 and 4, there is an analytic solution for the dipper function that expresses S_t in terms of S_m. This analytic expression is useful for fitting masking curves when there is both facilitation and Weber (unity slope) masking. The phase independence data of Beard et al.[23] argue against a single component masking model and would seem to imply a gain control mechanism. However, by introducing "energy or "complex cell" masking we are able account for the phase independence data using a saturating transducer framework. This is our Type 3 masking model.

A somewhat different approach that might lead to a similar two-component local gain is the data by Klein & Stromeyer[5] that provides evidence for inhibition between channels tuned to different phases. This study involved measuring the adaptation effect of first- plus third-harmonic gratings with different static and dynamic phase relationships. Simple gain-control mechanisms didn't account for the data, whereas phase inhibition seemed like a viable model.

Heeger's[2] article on gain control is "must" reading for insight into the power and remaining problems of contrast gain control as applied to cat cortical cells. Of importance to the present topic is that Heeger, from the beginning, uses an energy (Pythagorean sum of even and odd activity) formalism rather than a single component formalism. Our Type 3 model is a subset of Heeger's energy model, in that in our model the sum in the denominator extends over only a single energy term, that which appears in the numerator. One must be careful about applying an energy model to both numerator and denominator. We believe that to a first approximation it is fine to have the denominator summate over phase. However, in the numerator one should sum over phase only if phase uncertainty is clearly present. This would occur at medium to high spatial frequencies for narrow bandwidth stimuli. For broad bandwidth stimuli one can typically determine the phase, so the numerator should probably have a contribution only from the mechanism with optimal phase.

4. MULTIPLICATIVE NOISE.

In a signal detection framework the detectability of a test pattern is the ratio of the neural activity generated by the test pattern divided by the noise. Stromeyer & Klein[7] showed how a model with both an accelerating (rather than saturating) transducer function and multiplicative noise proportional to d' could quantitatively fit their contrast discrimination dipper function data. We agree with Legge, Kersten & Burgess[34] who argue that this type of masking will be difficult to discriminate from Types 1 and 2. The main difference between this category and the previous two is the focus on noise. Masking Types 1 and 2 don't introduce noise explicitly and thereby assume the output noise is always a constant.

We think that measuring the ROC slope could clarify the issue. A flatter-than-unity ROC slope is associated with multiplicative noise. Unfortunately, uncertainty masking, as will be discussed in connection with Type 5 masking, can also give flat ROC curves, so the ROC slope isn't an unambiguous test. A shallow ROC slope can eliminate the simplest versions of Types 1 - 3 masking that shouldn't give shallow ROC slopes. Multiplicative noise and shallow ROC slopes can arise by a subject's fluctuating attention which leads to fluctuating thresholds. Fluctuating thresholds smears out the signal distribution while leaving the noise distribution alone. If one maintains a very strict criterion, one minimizes the effect of smearing of the distribution to lower levels of channel activity. However, if one uses a two-alternative forced-choice paradigm (2AFC) then the effective criterion is quite loose and therefore the estimated d' is strongly affected by the smeared signal distribution.

Another class of experiments that can distinguish the present type of masking from the previous two is contrast matching. Consider the masking task of Chubb & Sperling[35] and Cannon and Fullenkamp[36]. They find that the presence of surrounding masking stimuli can reduce the perceived contrast of a test patch, as would be expected from Type 2 masking (gain control). However, it is important to note that this reduction of perceived contrast is relatively small, typically less than 30%. A much larger change in matched contrast would be predicted by the contrast gain formalism. When the test patch is in the fovea this small reduction may be about the same as the reduction of detection threshold, but in peripheral vision the detection threshold can be elevated by a much greater amount. Type 4 masking is important because it can account for a large increase in detection threshold without much of a decrease in perceived contrast.

5. MASKING BY HIGH SPATIAL FREQUENCY BEATS.

One of the most perplexing examples of masking is the masking by beats[37,38] in which a pattern of beats produced by an amplitude modulated 8 + 10 + 12 c/deg grating masks a 2 c/deg grating. The beating pattern acts as if it has been put through a rectifying nonlinearity[39]. The opposite type of nonlinearity (saturating) would have been expected. It is possible that rectification-complex cell models such as proposed by Wilkinson, Wilson & Ellemberg[40] could account for the data. However, we suspect that it will be difficult for such a simple model to capture the full flavor of the data.

This type of masking has more generality than is generally appreciated. It is found in two types of image sampling. Consider first the blocking artifacts found in JPEG and other compression methods. For severe compression the luminances in each block are replaced by the mean luminance of that block. Harmon & Julesz' picture of Lincoln is a familiar example. The high spatial frequencies are concentrated at the block edges. Randomizing the phases of these high frequencies greatly reduces the masking of the low spatial frequency information.

A second example occurs when a one-dimensional image is sampled and each sample is displayed as a line[41,42]) or a point[43]. Klein & Levi (in preparation) found that thresholds are elevated once the samples are spaced by more than 2 - 3 min of arc. Rowan & Banks (in preparation) find that the 2 - 3 min spacing produces a threshold elevation even in peripheral, infant and amblyopic vision for which the 2 - 3 min sampling grid is invisible. This result implies that the threshold elevation is occurring early in the visual pathway, before the spatial filters can do their smoothing (probably before the ganglion cell stage).

We include these dramatic effects as examples of where our present vision models underestimate the amount of masking. Further experiments are needed to clarify the underlying mechanisms of this type of masking.

6. STIMULUS UNCERTAINTY.

Uncertainty effects are usually discussed in the context of detecting a pattern on a uniform field[44]. When a suprathreshold masking pattern is also present the uncertainty effects can get magnified. When the masker pattern gets confused with the test large threshold elevations can result. This category of threshold elevation may be the most important (it surely is in real world scenes), and it has been the least discussed, at least until this SPIE meeting. At this meeting, to our surprise, all of the five talks in our session were on the topic of how noise intrusion could elevate thresholds. Noise intrusion is so important that we will devote two categories to it. Type 7 masking is true noise intrusion that degrades even the stimulus-known-exactly ideal observer. Type 6 masking, on the other hand, occurs when an ideal observer wouldn't get the mask confused with the test. The real observer gets them mixed up because of uncertainty about the signal. This might be the most important type of masking, but there is sufficient confusion about how to deal with uncertainty effects that progress is slow in this area. To help clarify some of the issues we will ask two questions: a) what are some of the causes of uncertainty, and b) what decision strategies can be used to cope with uncertainty.

6.1. Causes of uncertainty.

6.1.1 Imperfect memory. One of the simplest and rarely discussed factors is poor memory. If one forgets (or never knew) what the stimulation due to the masker looks like then it could get confused with the stimulation due to the test. For example, in a contrast discrimination task suppose the observer can't remember the reference contrast. This case can be explored by having the two displays (mask and mask + test) presented simultaneously vs. successively. We have carried out two experiments that measure the importance of memory in a contrast discrimination task.

Hu, Klein & Carney[21] measured contrast discrimination of a pair of static sinusoidal gratings as a function of the spatial gap between the pair. For the three conditions of 1.2 min separation, 12 min separation and no simultaneous reference the jnd contrast thresholds were 3, 4 and 6.5 percent. For the nearly abutting case (1.2 min) the contrast discrimination could be achieved by an orthogonal mechanism detecting the luminance jump and minimal contrast masking was found. These data show that memory is pretty good, but does produce about a 60% loss of threshold compared to the 12 min gap.

Carney & Klein[20] did a memory test for two types of tasks: contrast discrimination and resolution. Two types of presentations were examined. In one case a line pedestal (mask) was presented across the entire screen and the test stimulus (a test line for contrast discrimination and a test quadrupole for resolution) was added to half the line. In the other case the pedestal was presented only to the half with the test. In the first case there was a simultaneous reference. For the resolution task the presence of the reference did not reduce thresholds at all, i.e. the stimulus shape could be memorized well. For contrast discrimination the simultaneous reference reduced thresholds by about 40%. Thus it is seen that in the case of contrast discrimination a substantial portion of the masking is due to memory problems. Contrast is harder to memorize than shape. Even with the simultaneous case there can be a memory problem since one has to switch one's attention from one side of the display to the other. During the time of the switch there could be a memory loss. One can easily think of stimuli that are much harder to memorize[45], where the memory loss would be a major factor in the masking. Klein & Tyler[46] developed a formalism based on higher-order autocorrelation functions that seemed to be useful for classifying harmonically related sinusoids according to the ease with which phase discrimination is possible. We believe that our classification scheme is also useful for specifying how easily a pattern's phase can be memorized. Higher-order patterns are much more difficult to memorize than lower-order patterns.

6.1.2. Hard-wired mechanisms or fixed attention pool. Paying attention to irrelevant information may be caused by hard-wired mechanisms. Complex cells and collator mechanisms are often discussed in this context. Their role as summators will be considered in Section 6.2.2, but their contribution would reduce sensitivity only if local or simple-cell mechanisms were unavailable to mediate detection. Also in this category is a fixed attention pool. It is possible that in peripheral vision there is a limit on how locally the attention mechanism can be focused; that may be set by the size of one hypercolumn, for example.

6.1.3. Forced attention due to transients. When the mask and test are presented synchronously and briefly (≈30 msec)[12], then an alerting mechanism might attract attention to the full mask even if the mask is much broader in position, orientation and spatial frequency than the test. This can produce a strong increase in stimulus uncertainty.

6.1.4. Identity confusion in peripheral vision. If the mask and test contain similar looking features it is easy for the subject to get the masker mechanisms confused with the test mechanisms. For example, in peripheral vision suppose the mask is separated from the test pattern but is similar in appearance. In the periphery, localization is poor and the observer is unable to attend to just the area with the test. The experiments of He, Cavanagh and Intrilligator[47] fall in this category. The presence of the mask will produce a high false alarm rate which degrades performance.

One strategy for testing stimulus uncertainty is to manipulate the appearance of the test and the mask so that they look different. For example, if the test and mask are separated spatially, put a red filter over the test region and a green filter over the mask region to disambiguate them. Strong evidence that disambiguation reduces masking in peripheral vision was found by Kooi et al.[48] using contrast polarity, shape, depth, color, eye of origin and contrast as cues. If the strong masking in peripheral vision had been due to lateral gain control, it is unlikely that it would be specific to all the stimulus attributes that were used as cues. The dramatic unmasking that occurred when the test was disambiguated implies that attentional factors play an important role in masking.

6.1.5. Polarity-specific Crawford masking. Carney, Klein & Hu[49] presented data on the detection of a thin line masked by a pattern of similar spatial shape. The test pattern was presented as a single flash. The mask was a long duration stimulus with sudden onset. When the test pattern was presented close to the mask onset, we found striking polarity specific effects. There was strong masking when a same polarity test pattern preceded the mask onset, and when an opposite polarity test just followed the onset. Bowen[50] found a similar effect. A possible explanation of this masking was a Type 6 (stimulus uncertainty) effect. A same-polarity test pattern just preceding the mask would appear as a mask that came on a bit earlier. Since the observer is uncertain about the exact timing of the mask, the test pattern would be invisible. A similar explanation might account for the opposite-polarity test that followed the mask. This explanation sounds good in theory, but in practice the polarity-specific masking extends for a longer time period than this explanation warrants. So further work is needed to account for this type of masking.

6.2. Decision rules for uncertainty and their consequences. The presence of uncertainty means that the observer is weighting the multiple channels non-optimally. An extreme example would have the observer attending to totally irrelevant channels that are not stimulated by the test pattern. We will discuss four decision rules for dealing with uncertainty: the ideal observer, summation, maximum output, and phase uncertainty. Klein[51] provides much greater mathematical detail concerning the first three of these rules.

6.2.1. Ideal observer. Consider first the two channel case in which one channel is the optimal channel for detecting the test pattern and the other channel is statistically orthogonal to the test. The mask can be assumed to add noise to both channels. For this case it is possible to give an elegant analysis of uncertainty and to compare the ideal observer to various degraded observers[51,52]. It is found that as d' goes to zero the optimal decision rule is asymptotically a summation over all the attended mechanisms. This rule dominates when the stimulus-known-exactly d' is less than about unity. For larger values of d' the ideal observer rule gets quite close to a maximum output rule[53]. It is often stated that the ideal observer can be replaced by the maximum output rule. Not so. When the number of attended irrelevant channels is not high (say less than 10) and the d' is low, the maximum rule substantially underestimates low values of d'. As will be shown, the ideal observer has a linear transducer function in this regime, whereas the maximum rule is sharply accelerated.

6.2.2. Summation across mechanisms or complex cell pooling. Stimulus dilution. We have in mind something like a complex cell or the final pooling stage of Klein & Levi's "viewprint" model[6] that simply does a weighted sum of mechanism outputs. For low d' values this rule is optimal, but for high d' it grossly underestimates the visibility of the test pattern. This type of pooling would affect sensitivity only if local or simple-cell mechanisms were unavailable or of lower sensitivity for some reason. However, there are many reasons, covered in Section 6.1 for why the observer might be unable to attend to the proper mechanism, even when it has a high d'. Wilkinson, Wilson & Ellemberg[40] have carried out a number of experiments that lead them to believe in this type of hard-wired mechanism. In their experiments they tried to decrease the uncertainty by cueing the test stimulus with a pair of dark markers above and below the test. They found that the markers did not help. This result implies that attentional mechanisms do not have the machinery for good resolution in peripheral vision.

The predicted transducer function for the summation rule is simple to state. If a total of M channels are attended and only one has the signal, then the d' becomes:

$$d'_{1\text{-of-}M} = d'_{1\text{-of-}1}/M^{1/2} \tag{9}$$

where $d'_{1\text{-of-}1}$ is linearly related to the stimulus contrast. The linear behavior is because at very low contrasts the summation rule leaves the stimulus strength unchanged but increases the effective noise by a factor of $M^{1/2}$. This linear behavior at low contrasts may be surprising since one thinks of the 1-of-M transducer function as having acceleration. If M is very large then the transducer function is difficult to distinguish from a sharply accelerated function. But in most discrimination experiments the amount of uncertainty isn't large and this summation rule becomes important, and quite different from the less optimal maximum rule.

We should point out that we are being somewhat sloppy in calling the function relating d' to contrast the transducer function. The $d'_{1\text{-of-}1}$ function in Eq. 9 is the true transducer function. The $d'_{1\text{-of-}M}$ function that is produced in Eq. 9 because of uncertainty shouldn't really be called a transducer function since its shape is a statistical property of the decision process rather than the transduction process. We allow ourselves to be sloppy since it is commonplace.

The loss of visibility due to summation can be thought of as a stimulus dilution effect. In contrast gain control masking the non-optimal stimuli contribute to the divisive inhibitory pool by contributing to the denominator of Eq. 1. In uncertainty masking, the non-optimal stimuli add to the numerator and dilute the effectiveness of the stimulus.

6.2.3 Maximum output rule. For M=2 the ideal observer prediction is quite easy to calculate. However, for larger values of M it gets complicated. For that reason most investigators switch to a maximum rule whereby the detection judgment is based on the mechanism with the largest output. Nolte & Jaarsma[53] showed that the maximum rule is close to ideal (note, however, our caveat that it has problems for low values of d' and low values of M). For d'>1, to first order, the maximum rule simply subtracts a constant value from stimulus known exactly (linear) d'. The amount being subtracted increases slowly with M. When a mask is not present and the d' is high, the effect of uncertainty when using the maximum rule (or the ideal observer rule) is not dramatic. However, when a masker stimulates the attended channels, the activity acts as noise and can dramatically raise thresholds.

6.2.4 Phase uncertainty. When measuring the visibility of sinusoids or Gabor functions, phase uncertainty becomes an important consideration. This is especially true in peripheral vision and at medium and high spatial frequencies in foveal vision. Since the underlying mechanisms have a Gabor-like receptive field, small amounts of position uncertainty get translated into phase uncertainty even for broad bandwidth stimuli like thin lines. Thus phase uncertainty should play a central role in vision modeling. It is surprising that it has not been thoroughly studied. Klein[54] has recently completed a study of phase uncertainty. One of the goals was to find an good approximation for the phase uncertain transducer function. The transducer function for an ideal observer who knows the stimulus including its phase, is given by:

$$d'_{known} = c \tag{10}$$

where c is the stimulus strength in units such that the channel noise has unity standard deviation. The transducer function for a similar ideal observer who doesn't know the phase is to an excellent approximation given by:

$$d'_{unknown} = 1.894\ ((c/Th)^2 + \sigma^2)^{1/2} - \sigma) \tag{11}$$

where Th= 1.706 and σ = .683. The 1.894 factor (equal to $1/((1+\sigma^2)^{1/2} - \sigma)$ was chosen so that d' = 1 when c = Th (the threshold).

This transducer function has many interesting properties but we will focus on four:

a) For a wide range of stimulus strengths (0.6 < c/Th < 3) the difference between the phase-unknown and phase-known transducers is approximately: $d'_{unknown} - d'_{known} \approx - 0.65$. In that large range of stimulus strengths the phase unknown loss of d' corresponds to a 1-of-M uncertainty with M around 5 or 6. By "1 of M uncertainty" we mean the observer is attending to M orthogonal and equal channels, only one of which has the signal.

b) For very low signal strengths, c/Th < 0.3, Eq. 11 is approximately quadratic in c, and the effective value of M increases dramatically.

The following two points are relevant to all types of uncertainty related masking, not just phase uncertainty.

c) Uncertainty can have a powerful effect on reducing the errors in threshold estimates. The transducer function in Eq. 10 has unity slope and that in Eq. 11 has a slope that is steeper by a factor of about 1.7 for stimuli between threshold and twice threshold (a common range for doing experiments). This implies that the introduction of phase uncertainty allows thresholds to be estimated at a given precision with about 1/3 of the trials that would be needed if the phase were known. Introducing more uncertainty, such as temporal or positional uncertainty would steepen the transducer function even further and would counter intuitively increase the precision of the threshold estimates. Of course threshold would increase, but that might be okay under many conditions. In order to benefit from the additional uncertainty one must be careful that one doesn't also introduce threshold fluctuations that could make the psychometric slope shallower.

d) In the presence of masking noise, thresholds will increase in proportion to the noise strength. If the masker doesn't provide cues to reduce the uncertainty, then we would expect that the shape of the d' vs. test contrast curve should remain unchanged. Thus the finding of Legge, Kersten & Burgess[34], that noise linearizes the transducer function (which had been quadratic with no noise), is surprising. Their result might be due to the fact that the masker provided a temporal cue for the test stimulus, as discussed earlier. Further experiments are needed to understand whether a noise masker consistently reduces the transducer exponent. If it does, that would be strong evidence in favor of an intrinsic accelerating nonlinearity as opposed to an uncertainty effect.

7. INTRUSIVE NOISE.

Our final type of masking is the case where the masker is noise that directly affects the test mechanism. This case differs from Type 6 masking where the masker needn't have directly affected the test mechanism, but rather it has its effect because the observer treats the channels stimulated by the mask as if they were test channels. The task is to detect the test pattern in noise, where the noise contributes to the test mechanism. Type 6 masking was distinguished by having the noise not substantially overlap the test, so that the noise intrusion was produced by a cognitive uncertainty or confusion between the test and pedestal. In Type 7 masking, the intrusion is real. Type 7 masking also overlaps with Type 4 masking since the human observer is not ideal, but rather acts as if the noise is multiplied by a factor. Burgess[55] showed that, for a disk detection task human observers can be within 50% of the efficiency of an ideal observer. That is, the subjective (effective) noise variance is doubled from its true value. Since efficiency is measured in energy units, the 50% number means that the test contrast was sqrt(2) higher than what an ideal observer would need. It should be noted that to achieve these high efficiencies the observer was allowed unlimited viewing time. Under more controlled conditions with limited viewing, a 25% efficiency is regularly found (Eckstein, personal communication). A 25% efficiency means the effective noise energy is four times the actual noise energy (in contrast units this is a doubling of effective noise contrast). Two items are special about this type of masking: 1) It has been well studied by medical radiology researchers[55,56,57]. The medical imaging community has been active in this area for many years and there is a mountain of studies that are relevant to the task of developing fidelity metrics for complex scenes. 2) In light of item 1 it is surprising that there has been so little contact between the medical imaging community, the basic vision research community, and the image quality community.

ACKNOWLEDGMENTS

This work was supported by the National Eye Institute grants EY04776 to SAK, EY7890 to CWT and Air Force contract F49620-95-C-0018 to TC.

REFERENCES

[1] Albrecht, D. G. and Geisler, W. S. "Motion sensitivity and the contrast-response function of simple cells in the visual cortex" *Visual Neuroscience* , **7**, 531-546, 1991.

[2] Heeger, D. H. "Normalization of cell responses in cat striate cortex" *Visual Neuroscience*, **9**, 181-197, 1992.

[3]Legge G. E. and Foley, J. "Contrast masking in human vision" *J. Opt. Soc. Am* **70**, 1458 - 1471, 1980.

[4]Wilson, H. R. "A transducer function for threshold and suprathreshold human vision" , *Bio Cyber..* **38**, 171-178, 1980.

[5]Klein, S. A. and Stromeyer, C. F. "On inhibition between spatial frequency channels: Adaptation to complex gratings" *Vision Research*, **20**, 459-466, 1980.

[6]Klein, S. A. and Levi, D. M. "Hyperacuity thresholds of 1 sec: theoretical predictions and empirical validation." *J. Opt. Soc. Am.*, **A 2**, 1170 - 1190, 1985.

[7]Stromeyer, C. F. and Klein, S. A. "Spatial frequency channels in human vision as asymmetric (edge) mechanisms" *Vision Research* **14**, 1409 - 1420, 1974.

[8]Geisler, W. S. and Albrecht, D. G. "Cortical neurons: isolation of contrast gain control" *Vision Research*, **32**,1409-1410, 1992.

[9]Ohzawa, I, Sclar, G. and Freeman, R. D. "Contrast gain control in the cat visual cortex" *Nature* **298**, 266-268, 1982.

[10]Malik, J. and Perona, P. "Preattentive texture discrimination with early vision mechanisms" *J. Opt. Soc. Am* ., **A7**, 923 - 932, 1990.

[11]Wilson, H. R. and Humanski, R. A. "Spatial frequency adaptation and contrast gain control" *Vision Research* **33**, 1133-1149, 1993.

[12]Foley, J. M. "Human luminance pattern-vision mechanisms: masking experiments require a new model" *J. Opt. Soc. Am.* **A11**, 1710-1719, 1994.

[13]Thomas, J. P. & Olzak, L. A. "Is contrast gain control driven by local contrast energy" in press: *J. Opt. Soc. Am.* A Special Issue 1997.

[14]Wilson, H. R. and Kim, J. "Dynamics of a divisive gain control in human vision" submitted to: Vision Research 1997

[15]Barghout-Stein, L., Tyler, C. W. and Klein S. A. "Partitioning mechanisms of masking: contrast transducer versus divisive inhibition", *SPIE Proceeding*, current issue, 1997.

[16]Klein, S. A. "Fidelity metrics and the test-pedestal approach to spatial vision", *Computational Vision Based on Neurobiology*, Teri B. Lawton, Editor, *Proc. SPIE* **2054**, 142 - 154.

[17]Nachmias, J. & Sansbury, R. V. "Grating contrast: discrimination may be better than detection" *Vision Research* **14**, 1039-1042, 1974

[18]Stromeyer, C. F., Madsen, J. C., Kronauer, R. and Klein, S. A. "Opponent movement mechanisms in human vision" *J. Opt. Soc. Am.* **A1**, 876-883, 1984

[19]Klein, S. A., Casson, E. and Carney, T. Vernier acuity as line and dipole detection. *Vision Research* **30**, 1703-1719, 1990.

[20]Carney, T. and Klein, S. A. "Resolution acuity is better than vernier acuity" *Vision Research* **37**, 525-540, 1997.

[21]Hu, Q.M., Klein, S. A. and Carney, T. "Can sinusoidal vernier acuity be predicted by contrast discrimination" *Vision Research* **33**, 1241-1258, 1993.

[22]Carney, T. and Klein, S. A. "Bisection from the test-pedestal perspective" *Vision Research* to be submitted in 1997

[23]Beard, B Klein, S. A. and Carney, T. Motion thresholds can be predicted from contrast discrimination. In press: *J. Opt. Soc.Am.* A Special Issue on Control of Visual Sensitivity, 1997

[24]Klein, S. A. "Spatial vision models: Problems and successes" Chapt. in Spatial Vision in Humans and Robots, Cambridge Univ. Press, 9-32, 1993.

[25]McFarland, D. Mc Whirter, N. D., and McCarthy, M. D. (Eds.) Guinness Book of World Records 1991, Bantam Books: NewYork, 1991.

[26]Campbell, F. W. & Kulikowski, J. J. "Orientation selectivity of the human visual system" *J. Physiol. (London)* , **336**, 359-376, 1966.

[27]Levi, D. M., Harwerth, R. S. and Smith, E. L. "Humans deprived of normal binocular vision have binocular interactions tuned to size and orientation" *Science* , **206**, 852-854, 1979.

[28]Phillips, G. C. and Wilson, H. R. "Orientation bandwidths of spatial mechanisms measured by masking" *J. Opt. Soc. Am.* **A2**, 226-232, 1984.

[29]Ross, J., Speed, H. D. and Morgan, M. J. "The effects of adaptation and masking on incremental thresholds for contrast" *Vision Research* **33**, 2051 - 2056, 1993.

[30]Campbell, F. W. and Robson, J. G. "Application of Fourier analysis to the visibility of gratings" *J. Physiol. Lond.* **197**, 551-566, 1968.

[31]Hayhoe, M., Benimoff, N. I. and Hood, D. C. "The time course of multiplicative and subtractive adaptation process" *Vision Research*, **27**, 1981-1996, 1987.

[32]Crawford, B. H. "Visual adaptation in relation to brief conditioning stimuli" *Proc. Roy. Soc.* **B134**, 283-302, 1947.

[33]Kersten, D. "Spatial summation in visual noise" *Vision Research*, **24**, 1977-1990, 1984.

[34]Legge, G.E., Kersten, D. and Burgess, A. E. "Contrast discrimination in noise" *J. Opt. Soc. Am.*, A **4**, 392-404.

[35]Chubb, C., Sperling, G. and Solomon, J. A. "Texture interactions determine perceived contrast" *Proc. Nat. Acad. Sci. USA*, **86**, 9631-9635, 1989.

[36]Cannon, M. W. and Fullenkamp, S. C. "A model for inhibitory lateral interaction effects in perceived contrast" *Vision Research* **36**, 115-125, 1996.

[37]Henning, G. B., Hertz, B. G. and Broadbent, D. E. "Some experiments bearing on the hypothesis that the visual system analyzes spatial patterns in independent bands of spatial frequency." *Vision Research*, **16**, 887-898, 1975.

[38]Nachmias, J. and Rogowitz, B. E. "Masking by spatially-modulated gratings" *Vision Research* **23**, 1621-1629, 1983.

[39]Smallman, H. and Harris, J. "Nonlinear visual distortion: an effective expansive nonlinearity from asymmetry in ON and OFF pathways" *Invest. Ophthal. Vis. Sci.* **37**, S232, 1996.

[40]Wilkinson, F., Wilson, H. R. and Ellemberg, D. "Lateral interactions in peripherally-viewed texture arrays", in press: *J. Opt. Soc. Am.* A 1997

[41]Burr, D. C., Ross, J. and Morrone, C. "Local regulation of luminance gain" *Vision Research* **25**, 717-727, 1985.

[42]Rainville, S. J. M. and Kingdom, F. A. A. "The mechanisms for detecting compressively sampled gratings" submitted to: *Vision Research* 1997

[43]Mulligan, J. B. and MacLeod, D. I. A. "Visual sensitivity to spatially sampled modulation in human observers" *Vision Research*, **31**, 895-905, 1991.

[44]Pelli, D. G. "Uncertainty explains many aspects of visual contrast detection and discrimination" *J. Opt. Soc. Am.* **A2**, 1508-1532, 1985.

[45]Lawton, T. "Effect of phase structures on spatial phase discrimination" *Vision Research*, **24**, 139-48, 1984.

[46]Klein, S. A. and Tyler, C. W. "Phase discrimination of compound gratings: generalized autocorrelation analysis" *J. Opt. Soc. Am.* **A3**, 868-879, 1986.

[47]He, S., Cavanagh, P. and Intriligator, J. "Attentional resolution and the locus of visual awareness" *Nature* **383**, 334-337, 1996.

[48]Kooi, F. L., Toet, A., Tripathy, S. P. and Levi, D. M. "The effect of similarity and duration on spatial interactions in peripheral vision" *Spatial Vision*, **2**, 255-279, 1994.

[49]Carney, T., Klein, S. A. and Hu Q. J. "Visual masking near spatiotemporal edges" *SPIE: Human Vision Visual Processing and Digital Display VI*. Bernice E. Rogowitz and Jan P. Allenbach, Eds, 2657, 393-402, 1996.

[50]Bowen, R. W. "Isolation and interaction of ON and Off pathways in human vision: Pattern polarity effects on contrast discrimination" *Vision Research* **35**, 2479-2490, 1995.

[51]Klein , S. A. "Double judgment psychophysics: problems and solutions" *J. Opt. Soc. Am.*, A **2**, 1568 - 1585, 1985.

[52]Graham, N. V. S. Visual Pattern Analyzers. Oxford: Oxford Univ. Press. 1989.

[53]Nolte, L. W. and Jaarsma, D. "More on the detection of one of M orthogonal signals" *J. Acoust. Soc. Am.* 41, 497-505, 1967.

[54]Klein, S. A. "Ideal observer transducer functions for detecting phase unknown sinusoids", Submitted to Vision Res, 1997.

[55]Burgess, A. E. & Colbourne, B. "Visual signal detection. IV. Observer inconsistency" *JOSA.*, A5, 617- 627, 1994.

[56]Burgess, A. E. "Statistically defined backgrounds: performance of a modified nonprewhitening observer model" *J. Opt. Soc. Am.* **A11**, 1237 - 1242, 1994.

[57]Eckstein, M. P., Ahumada, A. J., Watson, A. B. "Visual signal detection in structured backgrounds. II Effects of contrast gain control, background variations and white noise" submitted to: *J. Opt. Soc. Am.* 1997.