1 INTRODUCTION

The proposed research is predicated on several basic assumptions about low level visual processing, all of which are common to standard models of visual texture perception. First, we assume that preliminary to extracting object boundaries, identities, locations, etc. from the retinal input, vision first applies a battery of fast, spatially parallel image transformations whose response images reflect "the amounts of various kinds of visual substances present in the image" (Adelson & Bergen, 1991). It is useful to imagine these substance-sensing transformations as being implemented in retinotopically organized neural arrays. In a given such array, all neurons are assumed to perform in parallel the same computation on the visual input impinging on a small region of the retina, but at different locations in the visual field. Thus any such "substance"-sensing array operates like a movie camera to make available to higher level vision a "neural image" (Robson, 1980) reflecting the rapidly changing distribution of a particular "substance" across the visual field. Because these hypothetical arrays are assumed to operate automatically, prior to conscious effort, we call them "preattentive mechanisms," or just "mechanisms." Each mechanism requires a vast, dedicated array of neurons, all of which are continuously active in normal vision. The high expense in neural computational resources of a single such mechanism makes it likely that human vision has only a modest number of them, raising the prospect that we may be able to catalogue them. Thus, two of the most compelling open questions in the field of low-level vision are: how many preattentive mechanisms exist in human vision, and what properties of the visual input do these mechanisms sense? Or, more grandly: what are the elementary substances of human vision?

It is important to note, however, that just because a mechanism is used for preattentive discrimination does not mean that it cannot also be brought under attentional control. Breathing takes no attention, yet we can bring our breath under attentional control to blow out a candle. We assume similarly that when we select a rock to trust our boot to as we hike, we are using the same mechanisms that preattentive vision uses, but bringing them under attentional control. To judge whether a rock face will hold our boot without slipping we must use what we shall call an "attentional filter" tailored to this particular task. An attentional filter takes spatially distributed image data as input and produces as output a neural image that can guide a specific sort of action. The attentional filter used in the hiking example aims to produce a "map" that reflects (as well as possible) the "grippiness" of the rock faces in view.

2 OVERVIEW OF THE PROPOSAL

2.1 Intellectual Merit

Perhaps the most fundamental question in vision science is: Which physical differences in the visual input are spontaneously visible and which are not? At present this question remains unanswered. The answer will take the form of a Table of the Dimensions of Preattentive Visual Sensitivity, or DPVS Table. Trying to theorize about vision without the DPVS Table is like trying to theorize about chemistry without the Periodic Table of Elements, and just as one could not fully have anticipated the consequences of Mendeleev's momentous discovery, it is difficult to foresee the theoretical and practical ramifications of the DPVS Table.

Some immediate benefits are clear, however. First, the DPVS Table will provide a firm foundation for the rapidly growing field of research into the visual perception of natural materials and surfaces. Indeed, our discovery of the "blackshot" mechanism is already influencing thinking in this field (e.g., Motoyoshi et al., 2007, Robilotto & Zaidi, 2006). In the absence of the DPVS Table, researchers in this field are shooting in the dark–latching onto ad-hoc image statistics that correlate

with human judgments; however, any predictive power these ad hoc statistics provide derives solely from their correlation to the dimensions of preattentive visual sensitivity catalogued in the Table.

More broadly, the DPVS Table is likely to prompt research into the brain systems that achieve the Table sensitivities. Such research may well lead to new theoretical insights, perhaps of clinical importance.

Goals The proposed research has three main goals:

1. We will derive the preattentive dimensions of visual sensitivity for three domains of physical variation.

2. In each of these domains we will also determine the spaces of attentional filters people can achieve for two very different visual tasks.

3. The results from 1 & 2 should enable us to

   (a) determine whether the achievable attentional filters are the same for the two tasks we use.

   (b) test the hypothesis that the attentional filters observers can achieve are linear combinations of the dimensions of preattentive visual sensitivity.

2.2 Broader Impact

The proposed research plan will directly involve a post-doc and a graduate student in active collaboration with the PI's and will indirectly involve many more undergraduates through roles in data collection. There will be regular meetings of all collaborators supported by the proposal. With only one graduate student and one post-doc there can be no guarantee that they will be women and/or under-represented minorities. That said, however, the record of the investigators is strong in this area. Each of Chubb and Sperling directs a lab at UCI that has successfully provided training at both the graduate and undergraduate level with a majority of students being women and under-represented minorities.

3 BACKGROUND

3.1 Mechanisms and discrimination

We assume that a target region is preattentively discriminated from a background only if the two regions differentially activate one or more mechanisms. For example, if they differentially activate one or more cone classes, target and background will be discriminable due to a difference in brightness or color.

In addition, however, since the work of Julesz (1962, 1975, 1981) and Beck (1966, 1982), it has been recognized that human vision has mechanisms sensitive to purely textural properties (see also Beck, Sutter & Ivry, 1987; Graham, 1989; Gurnsey & Browse, 1989; Julesz & Bergen, 1983). Many models of preattentive texture segregation have been offered (e.g., Beck, Prazdny, & Rosenfeld, 1983; Bergen & Landy, 1991; Bovik, Clark & Geisler, 1990; Caelli, 1985; Fogel & Sagi, 1989; Graham, 1989, 1991; Graham, Beck & Sutter, 1992; Grossberg & Mingolla, 1985; Knutsson & Granlund, 1983; Landy & Bergen, 1991; Malik & Perona, 1990; Wilson, 1993). All propose that human vision embodies a number of mechanisms sensitive to local pattern orientation and spatial frequency. Often the proposed mechanisms use spatially local linear filtering followed

by rectification. Julesz (1962) famously conjectured that all preattentive texture discrimination used linear filtering followed by a squaring nonlinearity. And indeed, spectral energy accounts well for many cases (Bergen & Adelson, 1988); however, many counterexamples exist (Diaconis & Freedman, 1981; Julesz, Gilbert, Shepp, & Frisch, 1973; Julesz, Gilbert, & Victor, 1978, Pollack, 1971a,b, 1972, 1973). Although such examples are suggestive, little progress has been made in filling in the DPVS Table (but see Victor & Conte, e.g., 1991, 1996).

4 PROPOSED RESEARCH

4.1 Overview

The proposed research uses three different psychophysical tasks to study preattentive sensitivity to physical variations in each of three domains. The three tasks are

1. the *discrimination task*, in which the observer views a briefly flashed field of texture and guesses the location of a target patch in a background of different texture.

2. the *comparison task*, in which the observer views a briefly flashed field comprising two patches of texture and has to use an attentional filter to judge which patch has the greater level of a given target property.

3. the *centroid task* in which the observer views a briefly flashed cloud of different sorts of elements and attempts to apply a target attentional filter to the cloud and extract the centroid of the attentionally filtered cloud.

The three domains of physical variation will be described more carefully below, but in brief they are:

1. *gray squares* (varying in Weber contrast).

2. *DoGs* (differences-of-Gaussians varying in contrast–see Fig. 4).

3. *colored squares* (equiluminant, and varying in saturation–see Sec. 4.6.3).

For a given domain of physical variation the three tasks reveal different things. The discrimination task is used to discover the dimensions of preattentive visual sensitivity that enable discrimination within that domain. By contrast, the comparison and centroid tasks are used to investigate the attentional filters the observer can achieve for two very different sorts of judgment.

The discrimination task requires the observer simply to detect the location of the target patch. This is possible only if his/her visual system has one or more mechanisms that are differentially activated by the target patch vs. the background. Sec. 4.4.1 explains how we use the discrimination task to (1) determine the number of mechanisms sensitive to variations in the domain under study, and (2) derive a basis spanning the space of sensitivity conferred by those mechanisms.

The comparison and centroid tasks (Secs. 4.3.3 and 4.3.4) both require the observer to try to achieve a target attentional filter. Secs. 4.4.2 and 4.4.3 explain how we can measure the attentional filter the observer actually uses to approximate the target filter in performing either task. By varying the target filter across different attentional conditions, we can derive the entire space of attentional filters achievable in each task.

The comparison and centroid tasks place very different demands on the observer. First, the comparison task uses densely textured displays of the same sort as are used in the discrimination task whereas the centroid task uses sparse displays of elements on a homogeneous background.

It seems likely that the dimensions of preattentive sensitivity to individuated elements differ in important ways from the dimensions of preattentive sensitivity to dense fields of the same sorts of elements. Second, extracting a centroid requires the observer to take a spatial average of the attentional filter output to derive an unknown location; by contrast, the comparison task requires the observer to take a difference between the attentional filter output across two regions with known locations. It may be that these different computational demands constrain the available attentional filters in different ways. Using both tasks should give us insight into these issues.

4.2 Previous research

We have made extensive use of both the discrimination (Sec. 4.3.2) and comparison (Sec. 4.3.3) tasks. We briefly review this work here.

Experiments using the discrimination task have shown that human vision has only three mechanisms that are differentially sensitive to textures consisting of randomly scrambled mixtures of different gray-levels (Chubb, Econopouly & Landy, 1994; Chubb, Landy & Econopouly, 2004; Chubb, Nam, Bindman & Sperling, 2007). Two of these mechanisms collectively confer sensitivity to the mean and variance of the gray-levels in the texture. Thus, human observers easily discriminate textures differing substantially either in their mean gray-levels (yielding a difference in brightness) or in the variances of their gray-level distributions (yielding a difference in "contrastiness"). As we discovered, however, human observers are also sensitive to third statistic, the "blackshot" of the texture, which is roughly the proportion of elements in the texture very close to black (Chubb, Econopouly & Landy, 1994; Chubb, Landy & Econopouly, 2004). Fig. 1 shows estimated blackshot sensitivity functions for 3 observers. However, even though we know (1) that human vision has

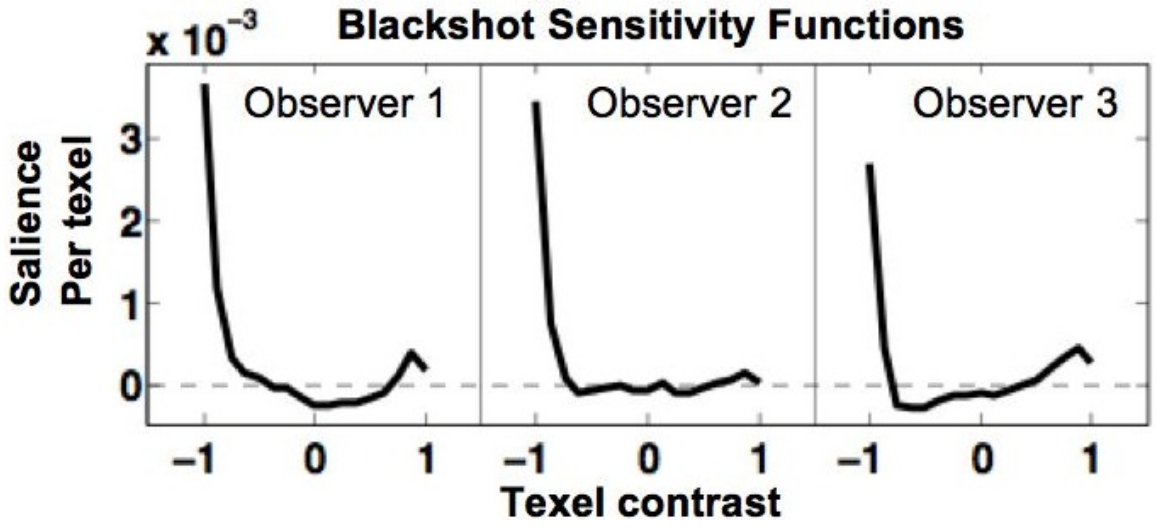

Figure 1: Blackshot sensitivity function estimates for three observers

two other mechanisms sensitive to these sorts of gray-level mixture textures and (2) what sorts of textures they collectively enable us to discriminate, we do not know the individual sensitivity functions of these two mechanisms.

We have also made a number of discoveries using variants of the comparison task (method first described by Chubb, 1999). First we (Nam & Chubb, 2000) discovered that when human observers are asked compare the means of mixtures of intensities, the attentional filter they achieve is linearly sensitive to texture element intensity (see Fig. 2), and hence very well matched to the attentional

task they were asked to perform. Note, however: this does not imply the existence in human vision

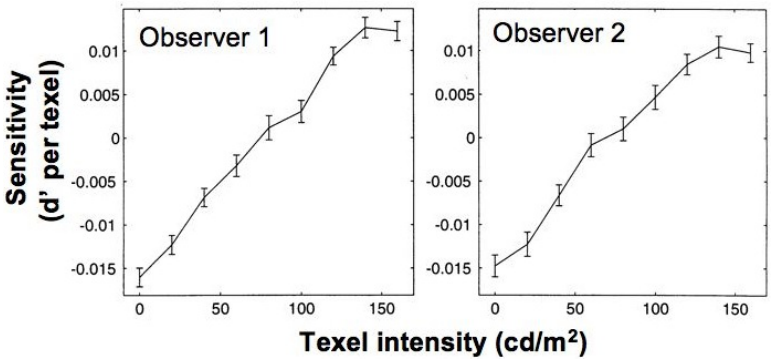

Figure 2: Attentional filters for mean of random mixtures of 9 different luminances

of a mechanism specifically tuned to texture mean; this finding implies only that human observers can construct an attentional filter well-tuned to texture mean from whatever mechanisms exist in their visual systems.

In another experiment, observers compared the variances of two texture patches (Chubb & Nam, 2000). Unlike the attentional filters observers achieved when comparing texture means, the attentional filters they achieved when comparing texture variances deviated substantially from the target filter. The ideal filter to detect texture variance is a parabolic function of luminance with a minimum at mean gray and maxima at black and white. However, as shown in Fig. 3 observers' attentional filters were dominated by texture elements darker than the mean. Despite receiving trial-by-trial feedback, observers behaved as though the task were to judge which texture had the greater total mass of dark elements, ignoring the bright ones.

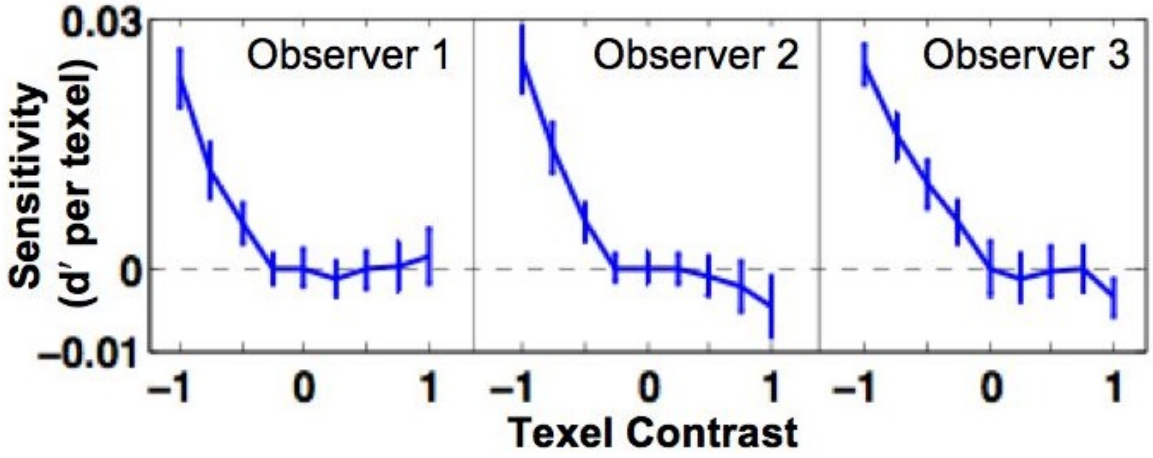

Figure 3: Attentional filters for variance of random mixtures of 9 different luminances.

In a variant of the comparison task paradigm, Chubb & Talevich (2002) studied the intensity-based attentional filters that observers could achieve in judging the predominant orientation of a field of texture. Observers judged the predominant orientation (up-&-right vs up-&-left) of a field of intermixed, oppositely oriented texture elements, attending selectively to some intensity-defined subset of elements. They found that observers can attend either to those elements darker than the mean (filtering out those brighter than the mean) or to those brighter than the mean (filtering out

the darker ones), or they can attend to all intensities, but they cannot achieve any other filters.

4.3 The tasks

Here we describe the discrimination, comparison and centroid tasks in detail. Both the discrimination and comparison tasks use textures called scrambles, described in the next section.

4.3.1 Scrambles

We always start with a particular set Ω of micropatterns. A micropattern is a "mini-image" that can serve as a single component in a large patch of texture. Some of the proposed experiments will use the micropattern set Ω_{DoG} is shown in Fig. 4a. The eight micropatterns in this set are all "Differences-of-Gaussians" (DoGs) of the same spatial form but varying in contrast polarity and amplitude. Other proposed experiments will use micropattern sets containing (a) gray squares of different Weber contrasts and (b) equiluminant squares of different saturations, and in some cases also of different hues. To generate a scramble from a micropattern set Ω we first specify

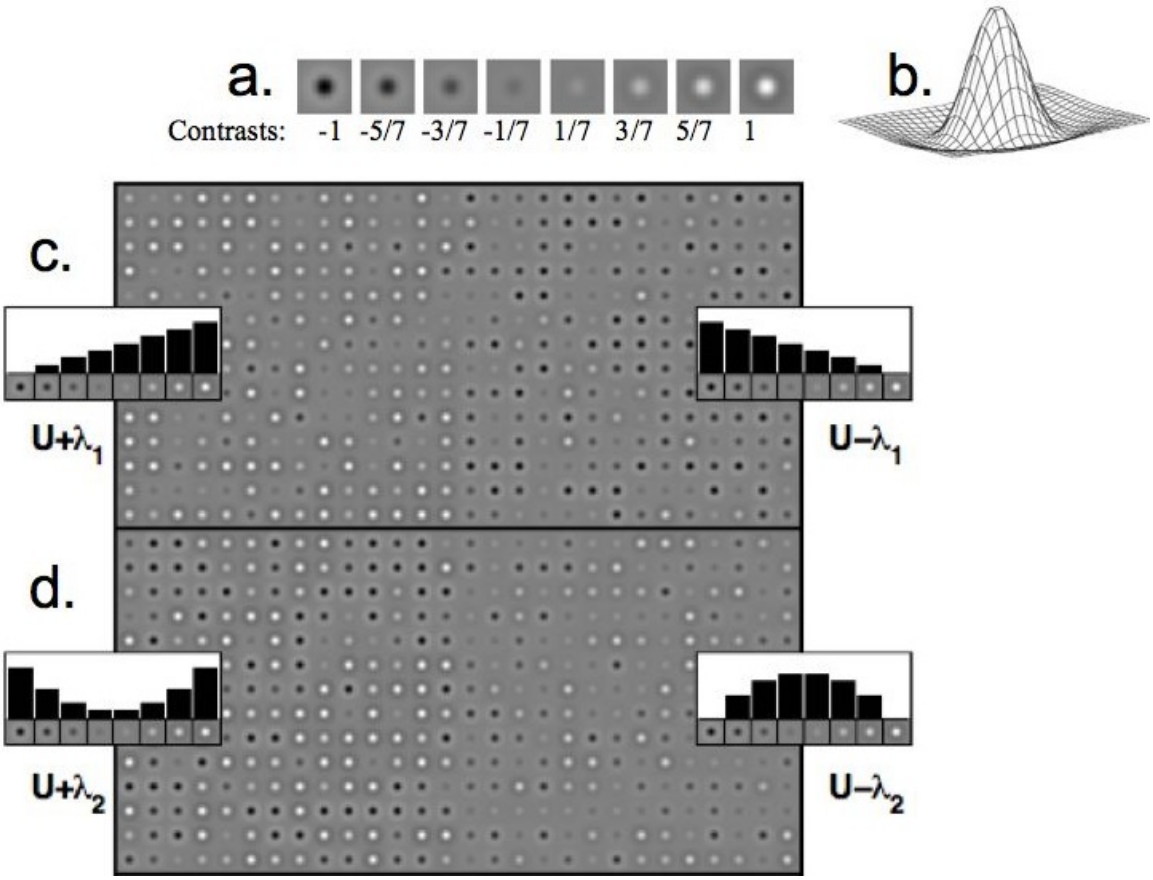

Figure 4: Scrambles of DoGs. a. The set Ω_{DoG} of micropatterns used, b. The profile of the micropattern with contrast = 1, c. Scrambles with histograms $U + \lambda_1$ (left) and $U - \lambda_1$ (right), d. Scrambles with histograms $U + \lambda_2$ (left) and $U - \lambda_2$ (right), for U the uniform histogram.

the proportions $p(\omega)$ with which different micropatterns $\omega \in \Omega$ appear in the scramble. Then we load a "virtual urn" with exactly the number of micropatterns needed to tile the stimulus

region in exactly (or as nearly as possible) the proportions $p(\omega)$; then we draw from the urn randomly without replacement to assign micropatterns to the texel locations of the scramble. We call probability distribution p the scramble histogram. Fig. 4c and d show the histograms of the component scramble patches on the left and right.

How many micropatterns should be included in Ω? We will typically include around $N_\Omega = 8$ different micropatterns (different levels of intensity) in Ω. The number 8 is a compromise. On the one hand, we need to be confident that N_Ω is greater than the number N of mechanisms sensitive to Ω-scrambles; otherwise we won't be able to determine N. On the other hand, if N_Ω is too large, then we lose the ability to significantly modulate the relative proportions of different micropatterns in a scramble. The fewer elements Ω contains, the easier it turns out to be to produce scrambles that that are easy to discriminate, and the more power one has to estimate sensitivity to different micropatterns in Ω.

4.3.2 The discrimination task

On every trial in the discrimination task the observer is asked to try to discriminate a target patch of Ω-scramble with some histogram from a background Ω-scramble with a different histogram. More specifically, for a number of reasons (discussed in Chubb, Econopouly & Landy, 1994), it is useful to use scrambles that are perturbed away from the uniform histogram $U(\omega) = \frac{1}{N_\Omega}$ in opposite directions (where N_Ω is the number of micropatterns in Ω). In other words, on every trial, we choose some perturbation function ϕ (called a modulator in our papers) and set the histograms of the target and background scrambles equal to $U + \frac{\phi}{2}$ and $U - \frac{\phi}{2}$ (which makes ϕ the difference between the two histograms).

We will use a 4AFC task in which observers view a blank field with a central fixation point and then are shown a stimulus for $200ms$. This stimulus will be a square patch composed of two patches of scramble, one with histogram $U + \frac{\phi}{2}$ and the other with histogram $U - \frac{\phi}{2}$, where perturbation ϕ is varied from trial to trial. One of the of the four possible stimulus configurations is shown in Fig. 5a. On each trial, one patch will take up either the top, left, right or bottom 3/8 of the stimulus and the other patch the complementary 5/8 of the stimulus. Thus the task is to judge whether the scramble boundary is above, to the right, below, or to the left of fixation. (This stimulus configuration has worked well in previous studies.)

As discussed below (Sec. 4.4.1), by choosing histogram differences ϕ appropriately, one can derive a basis spanning the space of all physical variations in a given domain to which people are sensitive. We call this the sensitivity space of the domain.

4.3.3 The comparison task

On every trial in the comparison task the observer briefly views ($\approx$ 200 ms) a pair of scrambles with histograms $U + \frac{\phi}{2}$ and $U - \frac{\phi}{2}$ (Fig. 5b), where the perturbation ϕ is varied experimentally across trials. In a given attentional condition, the observer attempts to judge whether the right or left patch has the higher level of a given target property. (E.g., in Nam & Chubb, 2000, the target property was the scramble mean.) Trial-by-trial feedback is provided to enable the observer to achieve the best attentional filter he/she can for any given target property.

As discussed below, for any target property, by appropriately varying ϕ it is possible to derive the attentional filter used by observers to make the required comparisons, and by varying the target property across different attentional conditions, one can determine the space of all achievable, comparison-task attentional filters for a given domain of physical variation.

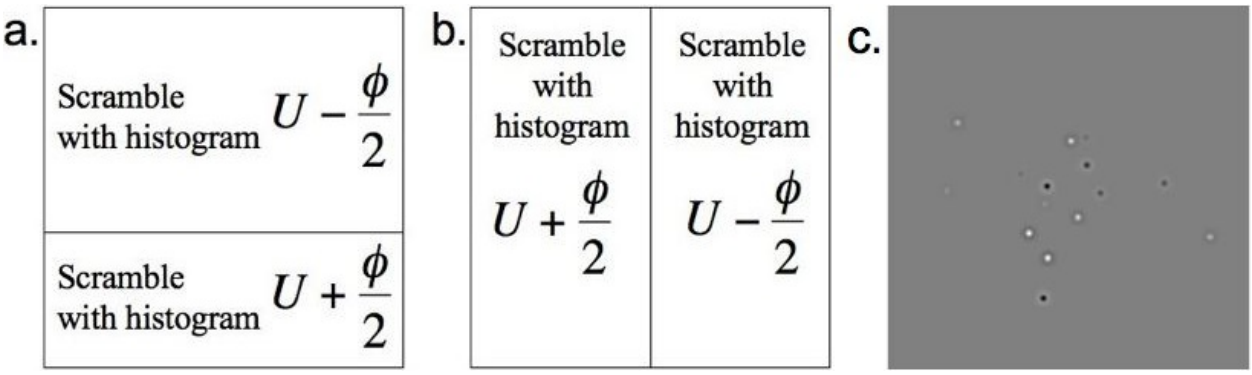

Figure 5: a. Stimulus configuration for discrimination task–Observer must detect the location of the boundary, which is below fixation in this display; it can also be above, to the right or to the left of fixation; b. Stimulus configuration for comparison task–the two regions to be compared are always on the left and right; the observer must judge which side has the higher level of the target property; c. 16-element dot cloud from centroid task using Ω_{DoG}–Given a brief display of this frame, observer must click on centroid of attentionally filtered cloud (e.g., centroid of the subcloud of DoGs with centers darker than the background).

4.3.4 The centroid task

On every trial in the centroid task the observer views a brief ($\approx$ 200 ms) display comprising a sparse cloud of elements drawn from one of our three micropattern sets Ω (either the gray squares, the DoGs or a set of colored squares). The cloud on a given trial will contain either one of each different type of element in Ω or two of each. Thus, for Ω_{DoG} (which has 8 micropatters) a given cloud will contain either 8 or 16 DoGs. (Fig. 5c shows a cloud with two of each.) Locations will be drawn from a circular bivariate Gaussian density with standard deviation appropriate to the type of micropattern used. In a given attentional condition, the participant attempts to apply a target attentional filter to the dots in the display and click on the centroid of the attentionally filtered stimulus. (For example, the observer might be asked to try to give equal weight to all DoGs with centers darker than the background and weight zero to all other DoGs.) After each trial the participant will receive feedback in the form of a 3-ring "target" centered on the correct response also showing a small marker at the location he/she clicked. We will conduct 400 trials (200 each of the 8-element and 16-element trials) in each attentional condition we test. (Our initial tests show that reasonable estimates of all model parameters can be derived from as few as 100 trials per each of the 8- and 16-element cloud-sizes. This paradigm seems remarkably powerful in the statistical sense.)

4.4 Methodology & Modeling

4.4.1 What we can learn using the discrimination task.

We use the discrimination task to derive the "sensitivity space" of a given domain Ω of physical variation. Here we first define "sensitivity space" and then describe how we measure it.

Mechanism sensitivity functions and the Ω sensitivity space. Think about any mechanism M in human vision. A basic assumption we make is that the average activation produced in M by a scramble is equal to the mean of the activations produced in M by the individual micropatterns

in the scramble (see Chubb, Landy & Econopouly, 2004, for discussion). If M responds more strongly to some micropatterns of Ω than to others, we say M is "differentially sensitive" to Ω. The function $f(\omega)$ that gives the response produced in M by an occurrence of ω in a scramble is called the "sensitivity function" of M. (f is analogous to a cone spectral absorption function, and micropatterns $\omega \in \Omega$ are analogous to different wavelengths of light.) Note that the difference in activation produced in M by an Ω-scramble with histogram $U + \frac{\phi}{2}$ vs. a background scramble with histogram $U - \frac{\phi}{2}$ is $f^T\phi$ (the inner product of f with ϕ). Let $f_1, f_2, \cdots, f_N$ be the sensitivity functions of all the mechanisms $M_1, M_2, \cdots, M_N$, in human vision that are differentially sensitive to Ω. Then we call the f_k's the sensitivity functions of Ω and the space they span the Ω sensitivity space.

Assumptions about how mechanisms combine to control discrimination. Suppose F is the (unknown) matrix whose column vectors are the Ω sensitivity functions. Then for any perturbation ϕ, the k^{th} entry of the vector $F^T\phi$ gives the difference in activation produced in the k^{th} mechanism by the scrambles with histograms $U+\frac{\phi}{2}$ vs $U-\frac{\phi}{2}$. We make the standard assumption that discrimination performance is a psychometric function of the Minkowski norm of $F^T\phi$ for some exponent β: i.e., of

$$Sal(\phi) = \left(\sum_k \left| f_k^T \phi \right|^\beta \right)^{\frac{1}{\beta}}, \tag{1}$$

where we think of $Sal(\phi)$ as the "salience" of the histogram difference ϕ. If $\beta = 2$, then the salience of a given histogram difference ϕ is the Euclidean length of $F^T\phi$, and equiperformance contours are parallel ellipses centered at 0. This case is particularly important because discrimination data tend to conform to such ellipses in practice. This has been observed for temporal contrast modulations (Rashbass, 1970; Watson & Nachmias, 1977), spatial contrast modulations (Logvinenko, 1993; Manahilov & Simpson, 2001), chromatic modulations (Knoblauch & Maloney, 1996; Noorlander & Koenderink, 1983; Poirson, Wandell, Varner & Brainard, 1990; Poirson & Wandell, 1996), and texture modulations (Victor, Chubb & Conte, 2004). We expect this to also be true in the proposed experiments. The discussion in Sec. 4.4.1 assumes that $\beta = 2$; however, the method we describe is robust with respect to broad violations of this assumption.

Deriving a basis of the Ω sensitivity space We now describe a method for deriving a basis of the Ω sensitivity space for an arbitrary micropattern set Ω. This method is new, remarkably powerful, and very general in scope of application. The method derives basis functions one at a time, with each successive element orthogonal to those previously derived. The method is essentially the same for each new element. We therefore describe the method for deriving the first element carefully and then summarize the iteration procedure.

Deriving the first Ω sensitivity space basis element Q_1. We start with an orthogonal basis $b_1, b_2, \cdots, b_{N_\Omega - 1}$ spanning the space of all histogram perturbations. (Note that there are only $N_\Omega - 1$ b_k's; this is because any constant function of Ω is useless for discrimination–thus, the Ω sensitivity space must be orthogonal to U.) We then pick a maximum amplitude perturbation ρ_{max} that yields superthreshold discrimination and measure the psychometric function $\Psi(A)$ that gives probability correct at discriminating histogram difference $A\rho_{max}$. We assume now that we can estimate the salience $Sal(\phi)$ of any histogram difference ϕ by

$$\widehat{Sal}(\phi) = \Psi^{-1}\left(\mathbf{P}(\phi)\right), \tag{2}$$

where $\mathbf{P}(\phi)$ gives the proportion correct at discriminating scrambles whose histograms differ by ϕ.

Next we show how to derive one function that resides in the Ω-sensitivity space. Let ρ be identical in form to ρ_{max} but scaled down to yield threshold discrimination. We will measure the form of the following real-valued function of Ω: $f_\rho = FF^T\rho$. Note that f_ρ is a linear combination of the columns of F and hence is an element of the sensitivity space, so we can use it in a basis of the sensitivity space. We call ρ the *seed* perturbation of f_ρ, and f_ρ the *expansion* of ρ.

To derive f_ρ we need only measure the inner products $f_\rho^T b_k$, $k = 1, \cdots, N_\Omega - 1$. We can then construct f_ρ as a linear combination of the b_k's. Assuming Eq. 1 holds with $\beta = 2$, it is easy to show that

$$Sal(\rho + \phi)^2 - Sal(\rho - \phi)^2 = 4f_\rho^T\phi \tag{3}$$

for any perturbation ϕ. Thus to estimate $f_\rho^T b_k$ for a given k, we measure performance at discriminating $\rho+\epsilon b_k$ and also at discriminating $\rho-\epsilon b_k$ (where ϵ is chosen small enough to keep performance in the steep part of the psychometric function). By applying Ψ^{-1} to the resulting discrimination probabilities, we obtain

$$f_\rho^T b_k \approx \frac{\left[\widehat{Sal}(\rho + \epsilon b_k)^2 - \widehat{Sal}(\rho - \epsilon b_k)^2\right]}{4\epsilon}. \tag{4}$$

We can then use the inner products $f_\rho^T b_k$ to derive our first basis function $Q_1 = f_\rho$. And a final point: it can also be shown that the estimate of $f_\rho^T b_k$ provided by Eq. 4 holds in the limit as $\epsilon \to 0$ regardless of the true value of β in Eq. 1, suggesting that our method is likely to be robust to failures of the assumption that $\beta = 2$.

Iterating the method to measure a full basis of the Ω-sensitivity space. We then derive a new basis of $N_\Omega - 2$ perturbations all orthogonal to Q_1, and reiterate the entire procedure to derive a second basis function Q_2 of the sensitivity space (orthogonal to Q_1). After some number N of iterations of this procedure, none of the next basis of perturbations (all orthogonal to all of Q_1, $Q_2, \cdots, Q_N$) will afford discrimination above chance. At this point we're done: The number of mechanisms differentially sensitive to Ω scrambles is N; Q_1, $Q_2, \cdots, Q_N$ make up an orthogonal basis of the sensitivity space; and the residual basis of indiscriminable perturbations spans the null space dual to the Ω sensitivity space.

4.4.2 What we can learn using the comparison task

The methods that use the comparison task (described in Chubb, 1999) have been applied several times (Nam & Chubb, 2000; Chubb & Nam, 2000; Chubb & Talevich, 2002). In brief: In a given attentional condition (using a given micropattern set Ω), the participant judges (with feedback) whether the left or the right scramble has the higher level of some target property P, fixed across that condition. The question is: what attentional filter f does the observer actually achieve in attempting to perform this task? Suppose the left-vs-right histogram differences tested empirically are ϕ_k, $k = 1, 2, \cdots, N$. Then on a trial with histogram difference ϕ_k, we assume the subject judges the scramble on the left higher in P than the scramble on the right if $f^T\phi$ plus noise is greater than 0. We assume the noise is normal. Thus, as long as our conditions are chosen so that the ϕ_k's span the space of all perturbations, we can use a generalized linear model with standard normal linking function to derive f.

To determine the space of all achievable attentional filters for the comparison task with a given micropattern set Ω one varies the target property P across different attentional conditions. We will use the following method (developed in Chubb & Talevich, 2002). Choose the first target property P_1 to make the comparison task as easy as possible; then derive the attentional filter f_1 used by

the observer to compare the levels of P_1 in different scrambles. Next choose target property P_2 (1) orthogonal to f_1 but (2) so that the comparison task is as easy as possible given (1); then derive the attentional filter f_2 the observer uses to compare the levels of P_2 in different scrambles. Reiterate this procedure, each time orthogonalizing target property P_k with respect to all attentional filters $f_1, \cdots, f_{k-1}$ previously achieved. After some number N of iterations, each new target property P_{N+1} tested will yield an attentional filter f_{N+1} that is well-approximated by a linear combination of the previously obtained attentional filters $f_1, \cdots, f_N$, and hence performance in comparing levels of P_{N+1} will be near chance. At this point, we're done: $f_1, \cdots, f_N$ are a basis of the space of achievable attentional filters.

4.4.3 What we can learn using the centroid task

Centroids and vision. Abundant evidence suggests that observers are good at visually estimating the centroids of spatially distributed targets (e.g., Morgan & Glennerster, 1991; McGowan et al., 1998; Zhou, et al., 2005). We use this skill as a lens through which to study attentional filters. To localize the center of a briefly viewed cloud of elements one must integrate spatial information that comes exclusively from preattentive mechanisms. Natural models of centroid extraction apply low-pass filters to the visual input and take appropriate maxima as centroid estimators. The primary question we address here is: what attentional filters can observers interpose between the visual input and the centroid filter?

Modeling. Fix all of (1) the micropattern set Ω, (2) the target attentional filter P, and (3) the number N of elements in the cloud. Then the data consist of the (x- and y-coordinates of the) locations clicked on by the observer in response to clouds presented on all trials. We fit a model with N_Ω parameters that minimizes the sum (across trials) of squared deviations of the predicted from the observed response locations. $N_\Omega - 1$ of these parameters are used to capture the attentional filter f used by the observer in the centroid computation (f uses only $N_\Omega - 1$ degrees of freedom because we require that $\sum_{\omega \in \Omega} |f(\omega)| = 1$). An unadulterated centroid computation is notoriously sensitive to peripheral elements in the cloud. For many purposes, it is useful to use "robust" estimators that discount the influence of such outliers. To discover whether our participants use robust estimators, we also include in our model a parameter α that captures any peripheral-vs-central bias. We call a bias in favor of central dots "robust" and a bias in favor of outlying dots "antirobust." Thus, the model we fit predicts responses by taking the centroids of the input clouds, where the ultimate weight assigned a cloud-element of type ω and location (x, y) is $f(\omega) \times DistanceCorrection_\alpha(x, y)$, where $DistanceCorrection_\alpha(x, y)$ is either less than or greater than 1 depending on the distance of (x, y) from the standard centroid derived using weights $f(\omega)$. The parameter α determines whether this correction is robust or antirobust.

4.5 New results.

The main purpose of this section is to prove that we can do what we claim we can. Previous research shows that we can use the comparison task to derive spaces of attentional filters (Nam & Chubb, 2000; Chubb & Nam, 2000; Chubb & Talevich, 2002). We therefore focus on the methods that use (1) the discrimination task to derive a basis of the Ω sensitivity space and (2) the centroid task to derive attentional filters and robustness/antirobustness parameters.

4.5.1 Deriving a basis of the Ω_{DoG} sensitivity space.

This section reports the results of our first application of the method described in Sec. 4.4.1.

Rationale. What mechanisms in human vision might be sensitive to Ω_{DoG}-scrambles? Consider a simple world in which (1) all sensitivity Ω_{DoG}-scrambles comes either from simple cells or complex cells, (2) simple cells are oriented, bandpass linear filters, and (3) complex cells gauge energy in a particular, oriented spatial frequency band. In this world, simple cells would be useless for discriminating Ω_{DoG}-scrambles because the space average response of any linear filter to any micropattern that integrates to 0 (true for all $\omega \in \Omega$) is 0. Thus, complex cells alone would be differentially sensitive to Ω_{DoG}-scrambles, and their sensitivity function would be parabolic with contrast (following the red curve in Fig. 6). Fig. 4 shows that this picture is wrong. Although the scrambles on the left and right in Fig. 4d have different contrast energy and would thus be discriminable by complex cells, the two scrambles in Fig. 4c have equal contrast energy, yet they are also easily discriminable. It was displays of this type that prompted Malik & Perona, 1990, to propose half-wave rectification as important in texture discrimination. The point is, however, that we don't have to speculate about the Ω_{DoG} sensitivity space; we can measure it.

Results thus far. We have used the new method described in Sec. 4.4.1 to measure two basis functions, Q_1 and Q_2, shown in the left and right panels of Fig. 6, of the Ω_{DoG} sensitivity space for one well-practiced observer. Also plotted in blue are the seed perturbations for each of Q_1 and Q_2. We chose the seed for Q_1 (a 6^{th} order polynomial) because pilot studies revealed that it was easily discriminable, yet its form was unexpected in light of standard theories, suggesting that its expansion might alter our understanding of low level vision. We emphasize, however, that for purposes of deriving a basis of the Ω_{DoG}-sensitivity space, any seed yielding threshold discrimination would have worked. Similar criteria were used to select the seed for Q_2.

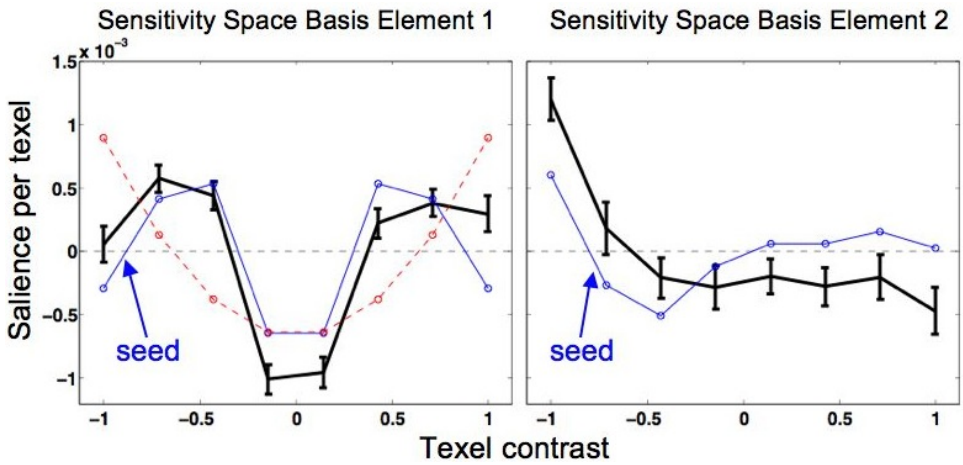

Figure 6: The first two basis functions Q_1 and Q_2 of the Ω_{DoG}-sensitivity space for one observer. The blue curves are the seed perturbations used to derive the two basis functions. Each seed yields threshold = 70% correct performance in the 4AFC discrimination task. The dashed red curve plots squared contrast–a sensitivity function predicted by many texture perception theories. Error bars give 95% CI's.

Discussion. Q_1 and Q_2 (Fig. 6) both show surprising sharpness in their tuning for DoG contrast. Many texture discrimination theories hypothesize mechanisms that use linear filters followed by squaring (texture energy mechanisms). Indeed, the Julesz conjecture (1962) held that all mechanisms had this form. Thus, one might have expected a dimension of sensitivity roughly matching

squared contrast (the red line in Fig. 6). Like this red curve, Q_1 is responds selectively to DoGs of high contrast of either polarity, but it saturates much more rapidly than expected from energy-based models.

In contrast to Q_1, Q_2 is sensitive to DoG contrast polarity, responding strongly to DoGs of negative contrast but not to DoGs of positive contrast. In this respect, Q_2 is similar to a negative half-wave rectifier. However (like the blackshot sensitivity function–Fig. 1), Q_2 saturates more rapidly than a standard half-wave rectifier.

The discovery of Q_1 and Q_2 represents an important step forward in our knowledge; neither was anticipated by previous theories. The adaptive roles of the sensitivities captured by these functions remain to be studied.

We have also observed that the space orthogonal to Q_1 and Q_2 contains discriminable perturbations, and we are in the process of extracting a third basis element of the Ω_{DoG} sensitivity space. We don't yet know whether more than three dimensions will be required.

4.5.2 Deriving a space of attentional filters for the centroid task.

This section reports results using the method described in Sec. 4.4.3 to analyze attentional filters available for localizing the centroids of clouds of micropatterns from Ω_{gray}, the set of small squares (0.2° in width) of Weber contrasts -1, -0.75, -0.5, -0.25, 0.25, 0.5, 0.75 and 1.

Method. On each trial the observer viewed the element-cloud for 200 ms on a background of mean luminance $34.2cd/m^2$. The cloud on a given trial contained either 8 elements (one of each $\omega \in \Omega_{gray}$) or 16 (two of each). The standard deviation of the cloud density was 5°. Responses were collected and feedback given as described in Sec. 4.3.4. Thus far we have tested 12 participants in each of three target attentional conditions:

1. *Attend-to-all*: Weight all dots equally, irrespective their different contrasts.
2. *Attend-to-dark*: Weight all dots darker than the background equally and ignore (i.e., give weight 0 to) dots lighter than the background.
3. *Attend-to-light*: Weight all dots lighter than the background equally and ignore dots darker than the background.

In each attentional condition, 100 trials were observed for each of the 8- and 16-element cloud-sizes.

Results. Mean attentional filters achieved are shown in Fig. 7. Participants were able to modify their attentional filters across conditions. The mean *Attend-to-all* filter (for each of $N = 8$ and 16) is strikingly flat across all contrasts, showing only a slight drop in sensitivity for the two Weber contrasts closest to 0. In addition, the *Attend-to-light* filters for both 8- and 16-element clouds show high selectivity for positive vs. negative Weber contrasts; and the reverse is true of *Attend-to-dark* filters. However, that the *Attend-to-light* and *Attend-to-dark* filters show systematic deviations from their ideal forms. Strikingly, the weights given to negative Weber contrasts by the *Attend-to-light* filters and positive contrasts by *Attend-to-dark* filters are all significantly greater than 0.

The parameter reflecting robustness vs. antirobustness was surprising (See Sec. 4.4.3). All but one participant showed either significant robustness or antirobustness in at least one of the six (3 target-filters $\times$ 2 dot-cloud sizes) conditions tested. Remarkably, six participants showed exclusively antirobust biases; of these six, two were significantly antirobust in 5 of 6 conditions. Three showed only robust biases; of these, two were significantly robust in 4 of the 6 conditions. Two other observers showed a mixture of robust and antirobust biases across conditions.

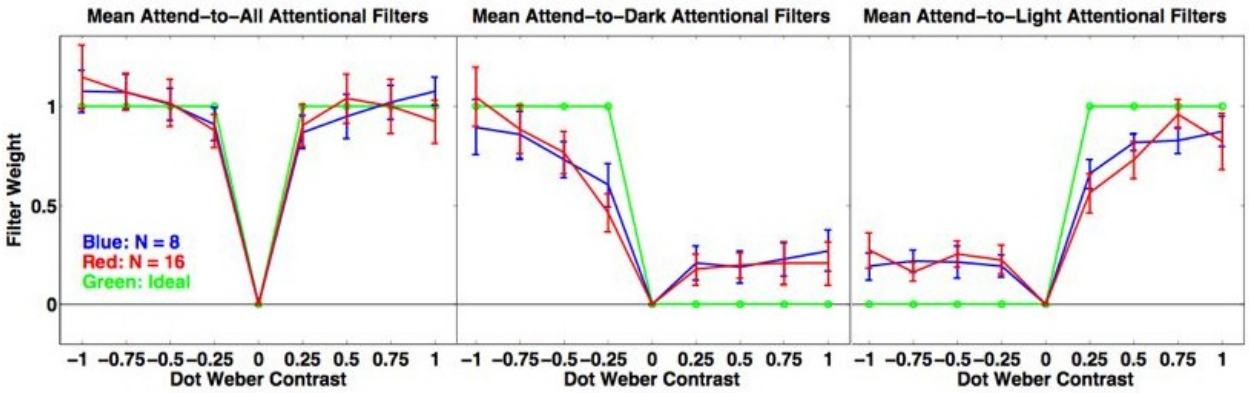

Figure 7: The mean attentional filters (of 12 participants) achieved in the centroid-extraction task. Blue (red) lines show filters for 8-dot (16-dot) clouds. Error bars give 95% confidence intervals.

Future work in this project. Fig. 7 shows that observers can achieve at least three different attentional filters in localizing centroids of Ω_{gray} clouds. Are any other filters achievable? To answer this question we will use additional attentional conditions with target filters orthogonal to those shown here.

4.5.3 Summary of proof-of-principle research.

We have described results to date in two experiments. One applies the method of Sec. 4.4.1 to derive the sensitivity space of Ω_{DoG}. The other explores the space of attentional filters achievable in localizing centroids of Ω_{gray} clouds. Both experiments, though still in progress, have already yielded important new results. All of the methods we use in the proposed experiments are guaranteed to yield unambiguous answers to fundamental questions of vision science.

4.6 Projected experiments

4.6.1 With Ω_{gray}

. Much of our past (and current) research has focused on deriving the Ω_{gray} sensitivity space and the attentional filters achievable for comparing Ω_{gray}-scrambles and also for extracting Ω_{gray} centroids. Our main goal in this domain is to complete the project of measuring the attentional filters achievable for both the centroid task and also the comparison task. We will also use the new method of Sec. 4.4.1 to confirm our previous estimates (obtained with different methods) of the Ω_{gray} sensitivity space.

4.6.2 With Ω_{DoG}

. We will use the method of Sec. 4.4.1 to finish deriving the Ω_{DoG} sensitivity space for the observer whose first two basis functions are shown in Fig. 6. We will also derive bases of the Ω_{DoG} sensitivity space for several other observers. If possible, for the same observers we will also derive the spaces of attentional filters achievable for comparing Ω_{DoG}-scrambles and also for extracting Ω_{DoG} centroids.

4.6.3 Investigations of preattentive chromatic sensitivity

Analyzing saturation in cardinal half-axes of DKL space. Here we select several hues, and for each hue, h, we take a set of small equiluminant squares of hue h, ranging in saturation from 0 (gray) to the maximum achievable saturation of h as our micropattern set. We will use 8 linearly increasing levels of saturation. Of primary interest are the gray-to-green, gray-to-red, gray-to-yellow and gray-to-blue half-cardinal-axes of DKL space in the equiluminant plane (Derrington, Krauskopf & Lennie, 1984). Call the 8-element micropattern sets drawn from these half-axes Ω_G, Ω_R, Ω_Y and Ω_B. On the one hand, current theories of chromatic mechanisms give us no reason to think that more than one mechanism is sensitive to saturation variations in any of these half-axes. On the other hand, pilot experiments suggest that observers can discriminate Ω_G-scrambles that differ either in mean saturation (in which case the scrambles differ in average greenness) or in saturation variance (in which case the scrambles differ in what might be described as saturation "contrastiness"). This suggests that the Ω_G sensitivity space has at least 2 dimensions. Whether the same is true for the other half-axses is an open question.

We will use the methods of Sec. 4.4.1 to measure the sensitivity spaces of Ω_G, Ω_R, Ω_Y and Ω_B. In addition, in each of these domains of physical variation, we will determine the spaces of attentional filters observers can achieve for both the comparison and centroid tasks.

Testing the opponency of chromatic mechanisms The classical theories of Hering (1964) and Hurvich & Jamieson (1957), and also neurophysiological results (e.g., Conway, 2001) suggest that color processing may be fundamentally opponent in nature. If so, then (for example) any mechanism sensitive to red will also be (with opposite sign) sensitive to green. On the other hand, there are reasons to suspect the existence of "half-axis" chromatic mechanisms. Some of these reasons are discussed in Beer & MacLeod, 2000, who use adaptation experiments to argue that there exist mechanisms sensitive only to half-axis variations of the luminance cardinal axis, a result supported by our discovery of the blackshot mechanism (which is sensitive only to variations along the negative contrast half-axis). However, thus far the evidence of half-axis mechanisms sensitive to variations in the equiluminant plane of DKL space is inconclusive.

To address this question, we will measure the sensitivity spaces of micropattern sets $\Omega_{R\&G}$ covering the full $L-M$ axis of DKL space and $\Omega_{B\&Y}$ covering the full $S-(L+M)$ axis. Let $N_{R\&G}$ be the number of mechanisms used to discriminate $\Omega_{R\&G}$-scrambles (i.e., the dimensionality of the $\Omega_{R\&G}$ sensitivity space). If all sensitivity to the half-axis scrambles of Ω_R and Ω_G comes from opponent mechanisms, then any mechanism sensitive to Ω_R-scrambles will also be sensitive to Ω_G-scrambles, implying that $N_R = N_G$; moreover, since no mechanism exists that is sensitive only to Ω_R-scrambles or only to Ω_G-scrambles, we should also have $N_{R\&G} = N_R = N_G$. Alternatively, at the other extreme, it may turn out that $N_{R\&G} = N_R + N_G$. This would mean that none of the mechanisms sensitive to Ω_R are also sensitive to Ω_G, i.e., that no red-sensitive or green-sensitive mechanisms are opponent. Finally, we might find that $max\{N_R, N_G\} < N_{R\&G} < N_R + N_G$, implying that the number of non-opponent mechanisms is $N_{R\&G} - max\{N_R, N_G\}$. Analogous experiments will be performed for blue vs. yellow. Finally, we will measure the attentional filters observers can achieve in the comparison and centroid tasks using $\Omega_{R\&G}$ and $\Omega_{B\&Y}$.