

Title: Bayesian Approaches To Color Vision

Short Title: Bayesian Color

David H. Brainard

Department of Psycholog

University of Pennsylvania

Room 302 C-Wing

3401 Walnut Street

Philadelphia, PA 19104

Phone: 215-573-7579

FAX: 215-746-6848

Email: brainard@psych.upenn.edu

DRAFT OF JULY 2008: To appear in The Cognitive Neurosciences IV.

NOTE: Permission to reproduce figures has been requested, will forward as received.

Acknowledgments: Supported by NIH RO1 EY10016. I thank S. Allred, J.

Nachmias, and B. Wandell for helpful comments on the manuscript.

Introduction

Visual perception is difficult. One pervasive reason for this difficulty is that image formation and sensory transduction lose information about the physical scene, so that many different scenes could have caused the same image data. Color vision presents an opportunity to understand how the brain copes with this information loss, because our understanding of the information loss and the scene parameters of perceptual interest are well-developed. In this sense, color provides a model system for developing and testing theories that may have more general applicability. Of course, color perception is an important aspect of our perceptual experience, and understanding how it arises is also interesting in its own right. This chapter provides an introduction to Bayesian modeling of human color vision, and reviews two lines of work where the approach has been fruitful.

Fundamentals of Color Vision

The visual system assigns a color to essentially all viewed objects. The information available about an object's color is carried by the spectrum of the light reflected from it, as illustrated by Figure 1A. This spectrum, which we refer to as the color signal, is specified by its power at each wavelength.

The color signal is given as the wavelength-by-wavelength product of the illuminant power and the object's surface reflectance function, where the latter specifies the fraction of incident light reflected from the object. Thus the color signal confounds object properties with those of the illuminant. To provide a representation of object reflectance that is stable across changes of illuminant, the visual system must process the color signal to separate the physical effects of illuminant and object surface.

The post-receptoral visual system does not have direct access to the color signal. Rather, this spectrum is encoded by the joint responses of the retinal mosaic of cone photoreceptors. There are three classes of cones, each characterized by a distinct spectral sensitivity (Figure 1B). These are often referred to as the L, M, and S cones.

Each individual cone codes information about light as a scalar quantity, the rate at which its photopigment is isomerized. This rate confounds the overall intensity of the color signal with its relative spectrum. Thus two physically distinct color signals can produce the same isomerization rate in all three classes of cones (Figure 1C). Moreover, there is at most one cone at each retinal location. To obtain even trichromatic information about the spectrum of the color signal, the visual system must combine information from cones at different retinal locations; sampling of

the image by the retinal mosaic confounds spatial and chromatic image structure.

This brief review illustrates a series of stages where information about object spectral properties is lost: the color signal confounds object reflectance with the spectral power distribution of the illuminant; the retina as a whole contains only three classes of univariate cones, so that the color signal's full spectrum is represented by at most three numbers; and at each retinal location there is only one cone. Each of these stages of information loss produces ambiguity about the scene being viewed.

How does the visual system resolve ambiguity to extract a perceptually useful representation of object color? Since the cone responses do not completely determine the reflectance properties of the object, some additional constraints must be imposed. Here we apply Bayesian analysis as a framework to express these constraints and develop models of color perception.

Bayesian Principles

Basic Ideas

Bayesian statistics provides a general formulation that allows image data to be combined with prior assumptions to provide a

reasonable estimate of the physical scene configuration. Both the information provided by the data and the prior assumptions are expressed in the common language of probability distributions. These two types of information are then combined via Bayes Rule to produce a third *posterior* distribution that expresses what is known about the scene. A specific estimate of the scene configuration can then be extracted from the posterior, for example by taking its mean.

An example serves to illustrate the key ideas. Imagine a toy universe containing only one wavelength of light and one spatial location. An illuminant impinging on a surface is specified by an intensity i and the object surface reflectance is specified by a single number s . We refer to these as the *scene parameters*, as fixing their values specifies the physical scene. The color signal here reduces to a single number, $c = i s$. If an eye with a single photoreceptor images the scene, we can model its response r by the equation $r = c + n$, where n is additive noise. We then ask, what do the *image data* r tell us about the values of the scene parameters? Within the framework of Bayesian analysis, this is given by the *likelihood*, written $p(r \mid i, s)$. For any scene parameters, the likelihood tells us how probable any observed response r is. Figure 2A illustrates the likelihood for our example, for the case $r = 1$. When the image data are held fixed, the likelihood is a function of the scene parameters.

Two features of the likelihood are worth note. First, some pairs (i, s) lead to higher likelihood than others. This means that the image data provide some information about the scene parameters. Second, the likelihood function has a ridge of equal values along the hyperbola $is = 1$. The fact that the likelihood is equal along this ridge indicates that the image data provide incomplete information: there are multiple scene configurations that the image data do not distinguish.

Because the image data do not uniquely determine the scene parameters, some other principle must be invoked to resolve the ambiguity. The prescription provided by the Bayesian approach is to specify the statistical properties of the scene parameters. In our toy universe, for example, it may be that not all illuminant intensities occur with equal probability. This fact can be expressed as a *prior* probability distribution over the scene parameters (Figure 2B). Here the prior has a ridge parallel to the s dimension, indicating that all values of s are equally likely. But along the i dimension, there is a concentration of illuminant probability in the vicinity of $i = 2$. The prior is a function of the same scene parameters as the likelihood.

Bayes Rule says that to combine the prior and likelihood into the posterior, all that is necessary is to multiply the two functions point-by-point and normalize the result so that the total probability is unity (Gelman, Carlin, Stern, & Rubin, 2004). The result of this operation for our example is illustrated by Figure 2C. The use of multiplication makes intuitive sense: the posterior has high values when both the prior and likelihood are high and goes to zero if either the prior or likelihood is zero. The posterior is often written in the form $p(i, s | r)$. In the case shown the posterior has an unambiguous peak, even though both the likelihood and prior are individually ambiguous about the scene parameters. Often the scene parameters are estimated by the maximum or as the mean of the posterior; other methods are also available (Blackwell & Girschick, 1954; Brainard & Freeman, 1997; Maloney, 2002).

Bayesian Models

Bayesian analysis provides a framework for generating models that may be applied to specific perceptual phenomena. The task of the modeler is to express the content of interest as a likelihood and prior, and then to link the resulting estimate of the scene parameters to perception. The framework is useful to the extent that it consistently generates models that describe, predict, and clarify empirical data.

The first part of Bayesian modeling is to specify the likelihood, which amounts to understanding the process by which scene parameters determine image data. For vision, the likelihood is in essence a description of how light flows through a scene to produce the retinal image, how the retinal image is sampled by the photoreceptor mosaic, and how precise the photoreceptor responses are. In the case of color these factors are well-understood (Wyszecki & Stiles, 1982; Kaiser & Boynton, 1996; Wandell, 1995), which means that generating the likelihood portion of a Bayesian model is straightforward.

The second part of Bayesian modeling is to specify the prior. In cases where the likelihood is well-constrained, the prior carries the critical content of the model. One way to obtain a prior is to examine statistical regularities in the natural environment and to choose a prior that captures these regularities. The idea is that evolution and development have optimized visual processing for the environment in which we operate, and that statistical regularities in the environment are likely to be deeply embedded in visual processing (see Helmholtz, 1910; Attneave, 1954; Shepard, 1987; Shepard, 1992; Adelson & Pentland, 1996; Mamassian, Landy, & Maloney, 2002; Weiss, Simoncelli, & Adelson, 2002; Geisler & Kersten, 2002; Maloney, 2002; also Knill & Richards, 1996; Rao, Olshausen, & Lewicki, 2002).

Often not enough is known about natural scene statistics to completely determine a prior. In this case, one strategy is to choose a parametric form for the prior that is broadly consistent with available measurements and then to use psychophysical data to constrain these parameters. We (Brainard et al., 2006; Brainard, Williams, & Hofer, 2008) and others (e.g. Mamassian & Landy, 2001; Stocker & Simoncelli, 2006) have found that this hybrid approach leads to effective models, and allows good progress to be made with the Bayesian approach even as we seek improved methods for measuring and specifying priors.

We defer to the discussion a number of important issues related to the selection of priors in Bayesian models of perception. Such discussion is more cogent in the context of the specific models reviewed in this chapter.

The final step in developing a Bayesian model of perception is to link the output of the Bayesian calculation, which is typically an estimate of scene parameters extracted from the posterior, to a measurement of human performance. Often the linking hypothesis involves the assumption that that goal of perception is to provide an explicit representation of one of the scene parameters. For example, color appearance is often taken to be a perceptual correlate of object surface

reflectance. Given this idea, a natural linking hypothesis for object color appearance is that two objects, viewed in different scenes, will have the same color appearance when a Bayesian estimation algorithm returns the same estimate of surface reflectance for each (Brainard, Kraft, & Longère, 2003).

Bayes and Color Constancy

In daily life it is common to refer to objects as having a color appearance: “the red apple”, “the blue house”, “the green car.” Although this is effortless, it is also remarkable. The color signal reflected from any object varies with the illumination, so that a stable percept must involve post-receptoral processing. The stability of object color appearance is called *color constancy*. Both color constancy and its close cousin lightness constancy have long been the target of experimental investigation (e.g. Helson & Jeffers, 1940; Burnham, Evans, & Newhall, 1957; McCann, McKee, & Taylor, 1976; Arend & Reeves, 1986; see Katz, 1935; Brainard, 2004; Gilchrist, 2006).

A few empirical generalizations are easy to state. First, the data confirm the introspective conclusion that human vision often exhibits excellent color constancy: object color appearance changes less than would be predicted by the corresponding change in reflected light (Arend &

Reeves, 1986; Brainard, Brunt, & Speigle, 1997; Brainard, 1998). Second, color constancy is not perfect: object color appearance does change somewhat with the illuminant. Third, factors other than the illuminant affect object color appearance. A particularly important factor is the surface reflectances of other nearby objects in the scene (Kraft & Brainard, 1999; Gilchrist, 2006).

Given that the above general facts are well-established, it seems clear that our goal should not be to generate further broad confirmations. Rather, we should seek models that accurately predict the color appearance of any object when it is viewed in an arbitrary scene. This is not an easy task, as a combinatorial explosion prevents enumeration and direct study of all possible scenes. It is thus critical to build models that not only fit extant data but that also embody principles that make it likely that they will generalize well (see Wyszecki & Stiles, 1982, pp. 584-586; also Krantz, 1968; Brainard & Wandell, 1992).

Different approaches to developing models of color appearance may be distinguished by the nature of the core principles they embody, and how these might enable generalization. Mechanistic models are constrained by abstracted properties of neurons in the visual pathways (Stiles, 1967). Examples of such abstractions include chromatic adaptation

(von Kries, 1902) and contrast coding (Wallach, 1948; Land, 1986). These models will generalize well if the mechanisms revealed by simple stimuli continue to operate unaltered for more complex images.

A second modeling approach tackles generalization directly and involves explicitly formulating and testing rules of combination. An example would be additive prediction of the action of the superposition of two scenes on an object's color appearance from measurements of the effect of each scene alone (Brainard & Wandell, 1992). A complement to this approach is to identify grouping and segmentation principles that allow the decomposition of complex scenes into separate and simpler regions, and then to study effects of context within such regions and to consider how they interact (Gilchrist et al., 1999; Adelson, 1999).

A third approach is *computational* (Marr, 1982; Barrow & Tenenbaum, 1978). Here the guiding hypothesis is that perceptual representations may be understood as biological approximations to well-defined information processing tasks. The use of Bayesian algorithms to model perception is an example of this approach. It will generalize well if the underlying hypothesis is correct: as one moves to richer and richer stimulus configurations what must be updated is the modeled likelihood and prior, but no new idea is required.

A Bayesian Model of Human Color Constancy

Here we outline a Bayesian model of color appearance. The first step is to define the likelihood function. This is based on the following imaging model (Brainard & Freeman, 1997; Brainard et al., 2006). A collection of flat matte objects is lit by a single spatially uniform and diffuse illuminant. The j^{th} object is characterized by its surface reflectance $s_j(\lambda)$, which specifies the fraction of incident light reflected at each wavelength λ . The illuminant is characterized by its spectral power distribution $i(\lambda)$, which provides light power at each wavelength. The color signal reflected from the j^{th} object to the eye at each wavelength is simply $c_j(\lambda) = i(\lambda) s_j(\lambda)$. For this example, we assume that the spatial scale of each object is large enough that we may neglect the fine structure of the interleaved retinal mosaic, so that the information obtained from the color signal reflected from each object is the isomerization rates of the L, M, and S cones. These are easily computed from $c_j(\lambda)$ and the known spectral sensitivities of each cone class.

The isomerization rates arising from a given $c_j(\lambda)$ will be perturbed by noise, which can be modeled as additive and normally distributed. Given the surface reflectances $s_j(\lambda)$ of N objects in the scene and the

illuminant spectral power distribution $i(\lambda)$, the likelihood may then be computed as a multivariate normal distribution with dimension $3N$, whose mean is given by the mean isomerization of the L, M, and S cones for each object and whose covariance matrix represents the additive noise. As with the simple example of Figure 2, this likelihood function represents an underdetermined estimation problem.

To specify a prior distribution for color constancy, we start with separate distributions over surface reflectance functions and illuminant spectral power distributions. The same general form is used for both, and is illustrated for surfaces in Figure 3. Following a number authors (Cohen, 1964; Buchsbaum, 1980; Maloney & Wandell, 1986; see also Maloney, 1986; Jaaskelainen, Parkkinen, & Toyooka, 1990) we first assume that each surface reflectance function $s_j(\lambda)$ may be approximated within a three-dimensional linear model. This means that we can write $s_j(\lambda) = w_j^1 s^1(\lambda) + w_j^2 s^2(\lambda) + w_j^3 s^3(\lambda)$, where the three *basis functions* $s^i(\lambda)$ are held fixed (see Figure 3A, B). The basis functions can be determined by applying principle components analysis to a large set of measured surface reflectance functions.

Within the three-dimensional model, each surface is specified by its triplet of weights w_j^1, w_j^2, w_j^3 . To induce a prior distribution over surfaces, we measure the three-dimensional histogram of weights corresponding to an ensemble of surfaces and fit this with a multivariate normal distribution (Figure 3C, Brainard & Freeman, 1997). Doing so allows us to compute the prior probability of any surface reflectance $s_j(\lambda)$. The same approach may be used to determine prior probabilities for illuminants $i(\lambda)$.

Delahunt and Brainard (2004b) conducted experiments in which observers were asked to judge the appearance of a test patch embedded in the simulated images of 17 scenes. Images of 4 of the scenes are shown in Figure 4A, with the location of the test patch indicated by the black rectangle in each image. Observers adjusted the chromaticity of the test patch until it appeared achromatic. Psychophysical measurements of this sort establish the chromaticity of the *achromatic locus* at the test patch location within each image, and this provides an excellent first order characterization of how scene context affects the color of any object (Speigle & Brainard, 1999). Across the set of scenes, both the illuminant and the surface reflectance of the background surface were varied. Figure 4B plots the chromaticity of the illuminant for each example scene (large

open circles) along with the achromatic loci (large filled circles). For three of the scenes, the chromaticity of the achromatic locus is near to that of the illuminant. These are the scenes where only the illuminant varied, and this pattern indicates good color constancy (see Delahunt & Brainard, 2004b; Brainard et al., 2006). For the fourth scene, where the background reflectance was also manipulated, the achromatic locus is far from the illuminant chromaticity; constancy was poor under this manipulation.

Brainard et al. (2006) used the general form of likelihood and prior described above, and implemented a variant of Brainard and Freeman's (1997) Bayesian algorithm that estimates the chromaticity of the scene illuminant from image data. They used a prior for surfaces that captured the structure of a large set of measured reflectances (as in Figure 3) and explored a parametric set of illuminant priors, all based on the CIE linear model for daylight. Brainard et al. (2006) then used the Bayesian algorithm to estimate the illuminant chromaticities for the same 17 scenes studied by Delahunt and Brainard (2004b). To link the illuminant estimates to the measured achromatic loci they interpreted the achromatic loci as representing the chromaticity of light reflected from a fixed achromatic surface. This allows mapping of estimated illuminant chromaticities to predictions of the achromatic loci (Brainard et al., 2006).

Figure 4B shows the predictions derived in this fashion (small open circles). The prediction for the leftmost image is guaranteed to be correct because this datum was used to infer the reflectance of the achromatic surface. The predictions for the other three scenes, and indeed for the entire collection of 17 images (see Brainard et al., 2006), are good. Of importance is that the agreement occurs both for cases where the observers showed good constancy and where constancy was poor.

Bayes and the Cone Mosaic

The treatment of color constancy above addresses the information loss caused by the interaction of surfaces and illuminants in the formation of the image, and the reduction from a full spectral representation of the color signal to the trichromatic representation provided by the L, M, and S cones. It assumed, however, that the responses of all three cone classes were available at each spatial location. This is intuitively reasonable if the spatial scale of the image of various objects is large compared to the spacing between cones. None-the-less, it would be more satisfying to have a theory of how information from cones is combined across space to provide a representation that is effectively trichromatic at such spatial scales. To develop such a theory, we can again turn to a Bayesian analysis.

As with the color constancy case, developing a Bayesian model requires specifying the likelihood and prior. We begin by deciding what scene parameters need to be estimated. Here we take these to be an *ideal image*, which we conceive of as the isomerization rates of the L, M, and S cones at each image location, before optical blur or retinal sampling. From the ideal image, we can compute the mean isomerization rate for each of the actual cones in the interleaved mosaic. The computation incorporates effects of blurring by the eye's optics as well as the location and type of each individual cone. We then use an approximation of additive normal noise to convert the mean isomerization to the likelihood function.

To develop an appropriate prior, we rely on a few well-established observations about natural images. First, within cone class images vary slowly over space. This is often expressed by the observation that the energy in natural images falls off as $1/f$, where f is spatial frequency (Field, 1987). Alternately, the same fact may be described as a strong correlation between image values at nearby locations within each cone class (Pratt, 1978). Second, at any given location there are high correlations between the isomerization rates of the L, M, and S cones (Burton & Moorehead, 1987; Ruderman, Cronin, & Chiao, 1998). This correlation occurs because typical color signals are spectrally broad-band and vary slowly with wavelength, because the spectral sensitivities of cones are

themselves broad, and because the spectral sensitivities of different cone classes overlap. We instantiate these strong spatial and spectral correlations in our prior by using a multivariate normal distribution whose covariance matrix is separable in space and color. This allows expression of the appropriate correlations (Brainard, 1994; Brainard, Williams, & Hofer, 2008). The use of a normal prior does not describe all of the regularities of natural images, but is strong enough to enable an effective algorithm.

Combining the likelihood and prior described above allows an estimate (Brainard, 1994; Brainard, Williams, & Hofer, 2008) of the ideal image from the isomerization rates of cones in the mosaic. Figure 5 illustrates estimator performance for a low spatial frequency stimulus. For such stimuli, the estimator returns near veridical estimates of the ideal image. This justifies the typical assumption that the fine structure of the interleaved mosaic may be neglected for most treatments of color vision.

It is possible to present stimuli that cause individuals with different cone mosaics have distinguishable perceptions. Hofer et al. (2005) used adaptive optics to present very small flashed monochromatic spots to observers whose chromatic topography had also been mapped. Figure 6 summarizes the results of these experiments. The top row (A) shows

mosaics for five individual observers. The second row (*B*) shows histograms of the aggregate naming behavior of each observer. There is large individual variation. Of particular note is that individuals with extreme L:M cone ratios tended to see more flashes as ‘white’ than individuals with more equal cone ratios.

We simulated the experiment of Hofer et al. and obtained reconstructed spots for each simulated flash. These were then mapped to color names through their chromaticities. Figure 6C shows the predicted naming histograms for each observer. The predictions capture the broad patterns of the data well. In particular, the fact that highly asymmetric mosaics lead to more flashes named ‘white’ is a salient feature of the predictions, just as it is with the data.

Figure 7 provides intuition about the model’s predictions. First consider the left hand panels (*A* and *C*). These show a mosaic consisting only of L cones, and the Bayesian estimate of the ideal image when only the central L cone is stimulated. A mosaic consisting of only L cones is completely monochromatic, and its responses provide no information about the relative spectral composition of the stimulus. When the Bayesian algorithm is applied to data provided by such a mosaic, it must rely on the prior to resolve the spectral ambiguity. The prior had a mean

of CIE daylight illuminant D65, and the resulting estimate is correspondingly a bluish white.

The right hand panels (*B* and *D*) show a second mosaic with the same spatial arrangement of cones. In this mosaic, however, the central L cone is surrounded by a set of M cones. We again simulated stimulating only the central L cone and obtained the algorithm's estimate. Here the resulting spot is red. This occurs because the M cones near to the central L cone add information about the spectral composition of the ideal image that was not available in the all L cone mosaic. These M cones have isomerization rates of zero, indicating that there is no middle wavelength light at their locations. Because the prior includes a specification of strong spatial correlations within each color plane, the zero M cone responses say that had an M cone been present at the central location it too would have had a small response. Putting this together with the large observed L cone response leads to an estimate of a spot with more power in the long wavelengths than in the middle wavelengths, and the result appears red.

Although the actual individual mosaics are more complex than the examples shown in Figure 7, the same intuitions still apply. In mosaics where there are large regions of homogeneous cone types, the reconstructed spots will have to rely more on the prior mean and will tend

to be ‘white’ rather than a more saturated color. The detailed modeling results shown in Figure 6 play this and related intuitions out in detail.

Discussion

The work reviewed here links color appearance to the performance of Bayesian algorithms. The first model (Brainard et al., 2006) accounts for both successes and failures of human color constancy. The second model (Brainard, Williams, & Hofer, 2008) explains how the visual system integrates information from individual cones in the retinal mosaic to provide a seamless trichromatic percept under normal viewing conditions. This model successfully predicts the color appearance of spots small enough to stimulate single cones.

The two models share important common features. First, the modeling begins with an analysis of information loss between the scene parameters and the responses of the cones. Second, the core of each model is the specification of a prior distribution over the scene parameters of interest. The prior acts to resolve the ambiguity about the scene parameters that remains after the image data have spoken. Given that the likelihood is well-constrained, it is the prior that provides the content of each model. In both cases, the general parametric form of the prior was determined through an analysis of naturally occurring scenes. The

particular priors used to model the data were then tuned by comparison of model predictions to the data (see Brainard et al., 2006; Brainard, Williams, & Hofer, 2008).

The Bayesian approach leads to parsimonious models that account for a performance across a wide range of conditions. Consider the color constancy example. Many studies of color constancy only vary the illumination, while the surfaces in the scene are held fixed (e.g. McCann, McKee, & Taylor, 1976; Arend & Reeves, 1986; Brainard & Wandell, 1992; Hansen, Walter, & Gegenfurtner, 2007). Under these conditions, many computational models can predict the generally good constancy that is observed (Buchsbaum, 1980; Land, 1986; Brainard & Freeman, 1997). When the surfaces in the scene are covaried with the illuminant, constancy is a more challenging computational problem that can differentiate between models (Maloney & Wandell, 1986; Brainard & Wandell, 1986; Brainard & Freeman, 1997; Kraft & Brainard, 1999). The model presented here accounts for performance when the illuminant is varied alone, and when both illuminant and scene surfaces are covaried. This feature emerged as a consequence of the interaction of a reasonable prior with the likelihood, as implemented through Bayes rule. A similar parsimony characterizes the account of small spot colors.

The Bayesian approach provides a clear generalization path. Both models presented here have to date undergone only preliminary tests against data, and it is likely that conditions can be found that produce model failures. For example, a few classic and more numerous recent studies of color and lightness constancy have sought to generalize to stimulus conditions beyond diffusely illuminated flat matte surfaces (Hochberg & Beck, 1954; Gilchrist, 1980; Nishida & Shinya, 1998; Bloj, Kersten, & Hurlbert, 1999; Yang & Maloney, 2001; Boyaci, Maloney, & Hersh, 2003; Boyaci, Doerschner, & Maloney, 2004; Doerschner, Boyaci, & Maloney, 2004; Delahunt & Brainard, 2004a; Ripamonti et al., 2004; Xiao & Brainard, 2008; Motoyoshi, Nishida, Sharan, & Adelson, 2007; also Fleming, Dror, & Adelson, 2003). It is not surprising that experiments of this sort reveal phenomena that cannot be accounted for by the Bayesian constancy model presented above, since that model is based on a likelihood and prior that do not allow for spatially varying illumination, objects and scenes that vary in three-dimensions, nor objects with geometrically complex surface reflectance functions. On the other hand, the Bayesian prescription tells us that to build more general models, it may be sufficient to generalize the likelihoods and priors. This is less daunting than developing models *de novo*. Recently, models that may be conceived as generalizing the likelihood function to account for directional illumination and variation in object pose have been reported

(Boyaci, Maloney, & Hersh, 2003; Bloj et al., 2004; Boyaci, Doerschner, & Maloney, 2006).

Although the Bayesian approach has the appealing features described above, when employed alone it will not provide a complete account of vision. First, the approach is silent about mechanism. The way that the models are implemented on a digital computer need not speak to how equivalent performance is achieved by the nervous system. The value of a Bayesian model in understanding mechanism is indirect. The model tells us what input-output relationship any mechanistic model must satisfy.

One common misconception about Bayesian models is that because the likelihood and prior are clearly separated in the formulation, this separation should be expected in the physiological representation. That this is not necessary is easily grasped if one considers that there is a simple linear receptive field interpretation of the Bayesian model of trichromatic image reconstruction. For any choice of prior distribution, the appropriate transformation from cone responses to the estimated ideal image may be accomplished by weighted sums that link cone responses to each estimated L, M, or S value (Brainard, Williams, & Hofer, 2008). In a neural implementation of this sort, the properties of the likelihood and

prior are jointly and implicitly encoded in the specific values of the weights; neither the likelihood nor the prior appear explicitly.

An oft-asked question that arises in the context of Bayesian models is “where did the priors come from?” If we view the role of the prior in a Bayesian model as specifying how the visual system resolves the ambiguity in the likelihood, then it becomes clear that this question is not really about the priors. Rather, the question is really “how did evolution and development shape the visual system so that it operates effectively in the natural environment?” This is an excellent question, but one that applies to any model that correctly describes stable adult performance. The current absence of an answer is not a weakness specific to the Bayesian approach. Indeed, one appeal of the Bayesian approach is that the prior is represented explicitly. Thus, the prior used in a successful model converts the raw experimental data to a form that may be compared to measurements of the statistics of the natural environment (see Brainard et al., 2006). When the derived prior matches the environmental statistics, the Bayesian model provides a quantitative link between behavioral performance and the statistics of the environment. When there is a mismatch between the derived prior and environmental statistics, the modeling emphasizes that our understanding remains incomplete.

Finally, we close by noting there are interesting and important factors that the Bayesian approach, as elaborated here, does not include. Because information processing consumes energy, for example, there may be tradeoffs between optimal estimation and efficient processing. A number of analyses interpret mechanisms of early color vision as efficient solutions to information transmission and representation (Buchsbaum & Gottschalk, 1983; Derrico & Buchsbaum, 1990; Atick, 1992; van Hateren, 1993; Ruderman, Cronin, & Chiao, 1998; Parraga, Brelstaff, Troscianko, & Moorehead, 1998; von der Twer & MacLeod, 2001; Parraga, Troscianko, & Tolhurst, 2002; Lee, Wachtler, & Sejnowski, 2002; Doi, Inui, Lee, Wachtler, & Sejnowski, 2003; Caywood, Willmore, & Tolhurst, 2004; Wachtler, Doi, Lee, & Sejnowski, 2007). Although this work shares with the Bayesian models presented here an emphasis on the statistics of the ensemble of natural scenes, it differs in its emphasis: the Bayesian models presented in this chapter focus on veridicality rather than efficiency of representation. Additional insight is likely to be obtained when both estimation and efficiency are considered jointly (see Abrams, Hillis, & Brainard, 2007; Hillis & Brainard, 2007).

References

- Abrams, A. B., Hillis, J. M., & Brainard, D. H. (2007). The relation between color discrimination and color constancy: when is optimal adaptation task dependent? *Neural Computation*, *19*, 2610-2637.
- Adelson, E. H. (1999). Lightness perception and lightness illusions. In M. Gazzaniga (Ed.), *The New Cognitive Neurosciences*, 2nd ed. (pp. 339-351). Cambridge, MA: MIT Press.
- Adelson, E. H., & Pentland, A. P. (1996). The perception of shading and reflectance. In D. Knill & W. Richards (Eds.), *Visual Perception: Computation and Psychophysics* (pp. 409-423). New York: Cambridge University Press.
- Arend, L. E., & Reeves, A. (1986). Simultaneous color constancy. *Journal of the Optical Society of America A*, *3*, 1743-1751.
- Atick, J. J. (1992). Could information theory provide an ecological theory of sensory processing. *Network-Computation in Neural Systems*, *3*, 213-251.
- Attneave, F. (1954). Some informational aspects of visual perception. *Psychological Review*, *61*, 183-193.
- Barrow, H. G., & Tenenbaum, J. M. (1978). Recovering intrinsic scene characteristics from images. In A. R. Hanson & E. M. Riseman (Eds.), *Computer Vision Systems*. New York: Academic Press.
- Blackwell, D., & Girschick, M. A. (1954). *Theory of Games and Statistical Decisions*. New York: Wiley.

- Bloj, A., Kersten, D., & Hurlbert, A. C. (1999). Perception of three-dimensional shape influences colour perception through mutual illumination. *Nature*, *402*, 877-879.
- Bloj, M., Ripamonti, C., Mitha, K., Greenwald, S., Hauck, R., & Brainard, D. H. (2004). An equivalent illuminant model for the effect of surface slant on perceived lightness. *Journal of Vision*, *4*, 735-746.
- Boyaci, H., Doerschner, K., & Maloney, L. T. (2004). Perceived surface color in binocularly viewed scenes with two light sources differing in chromaticity. *Journal of Vision*, *4*, 664-679.
- Boyaci, H., Doerschner, K., & Maloney, L. T. (2006). Cues to an equivalent lighting model. *Journal of Vision*, *6*, 106-118.
- Boyaci, H., Maloney, L. T., & Hersh, S. (2003). The effect of perceived surface orientation on perceived surface albedo in binocularly viewed scenes. *Journal of Vision*, *3*, 541-553.
- Brainard, D. H. (1994). *Bayesian method for reconstructing color images from trichromatic samples*. Paper presented at the IS&T 47th Annual Meeting, Rochester, NY, 375-379.
- Brainard, D. H. (1998). Color constancy in the nearly natural image. 2. achromatic loci. *Journal of the Optical Society of America A*, *15*, 307-325.

- Brainard, D. H. (2004). Color Constancy. In L. Chalupa & J. Werner (Eds.), *The Visual Neurosciences* (Vol. 1, pp. 948-961). Cambridge, MA, U.S.A.: MIT Press.
- Brainard, D. H., Brunt, W. A., & Speigle, J. M. (1997). Color constancy in the nearly natural image. 1. asymmetric matches. *Journal of the Optical Society of America A*, *14*, 2091-2110.
- Brainard, D. H., & Freeman, W. T. (1997). Bayesian color constancy. *Journal of the Optical Society of America A*, *14*, 1393-1411.
- Brainard, D. H., Kraft, J. M., & Longère, P. (2003). Color constancy: developing empirical tests of computational models. In R. Mausfeld & D. Heyer (Eds.), *Colour Perception: Mind and the Physical World* (pp. 307-334). Oxford: Oxford University Press.
- Brainard, D. H., Longere, P., Delahunt, P. B., Freeman, W. T., Kraft, J. M., & Xiao, B. (2006). Bayesian model of human color constancy. *Journal of Vision*, *6*, 1267-1281.
- Brainard, D. H., & Wandell, B. A. (1986). Analysis of the retinex theory of color vision. *Journal of the Optical Society of America A*, *3*, 1651-1661.
- Brainard, D. H., & Wandell, B. A. (1992). Asymmetric color-matching: how color appearance depends on the illuminant. *Journal of the Optical Society of America A*, *9*, 1433-1448.

- Brainard, D. H., Williams, D. R., & Hofer, H. (2008). Trichromatic reconstruction from the interleaved cone mosaic: Bayesian model and the color appearance of small spots. *Journal of Vision, in press*.
- Buchsbaum, G. (1980). A spatial processor model for object colour perception. *Journal of the Franklin Institute, 310*, 1-26.
- Buchsbaum, G., & Gottschalk, A. (1983). Trichromacy, opponent colours coding and optimum colour information transmission in the retina. *Proceedings of the Royal Society of London B, 220*, 89-113.
- Burnham, R. W., Evans, R. M., & Newhall, S. M. (1957). Prediction of color appearance with different adaptation illuminations. *Journal of the Optical Society of America, 47*, 35-42.
- Burton, G. J., & Moorehead, I. R. (1987). Color and spatial structure in natural images. *Applied Optics, 26*, 157-170.
- Caywood, M. S., Willmore, B., & Tolhurst, D. J. (2004). Independent components of color natural scenes resemble V1 neurons in their spatial and color tuning. *J. Neurophysiology, 91*, 2859-2873.
- Cohen, J. (1964). Dependency of the spectral reflectance curves of the Munsell color chips. *Psychon. Sci, 1*, 369-370.
- Delahunt, P. B., & Brainard, D. H. (2004a). Color constancy under changes in reflected illumination. *Journal of Vision, 4*, 764-778.

- Delahunt, P. B., & Brainard, D. H. (2004b). Does human color constancy incorporate the statistical regularity of natural daylight? *Journal of Vision*, 4, 57-81.
- Derrico, J. B., & Buchsbaum, G. (1990). A computational model of spatiochromatic image coding in early vision. *Journal of Visual Communication and Image Representation*, 2, 31-38.
- Doerschner, K., Boyaci, H., & Maloney, L. T. (2004). Human observers compensate for secondary illumination originating in nearby chromatic surfaces. *Journal of Vision*, 4, 92-105.
- Doi, E., Inui, T., Lee, T.-W., Wachtler, T., & Sejnowski, T. J. (2003). Spatiochromatic receptive field properties derived from information-theoretic analyses of cone mosaic responses to natural scenes. *Neural Computation*, 15, 397-417.
- Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A*, 4, 2379-2394.
- Fleming, R. W., Dror, R. O., & Adelson, E. H. (2003). Real-world illumination and the perception of surface reflectance properties. *Journal of Vision*, 3, 347-368.
- Geisler, W. S., & Kersten, D. (2002). Illusions, perception, and Bayes. *Nature Neuroscience*, 5, 508-510.

- Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). *Bayesian Data Analysis* (2cd ed.). Boca Raton: Chapman & Hall/CRC.
- Gilchrist, A. (2006). *Seeing Black and White*. Oxford: Oxford University Press.
- Gilchrist, A., Kossyfidis, C., Bonato, F., Agostini, T., Cataliotti, J., Li, X., Spehar, B., Annan, V., & Economou, E. (1999). An anchoring theory of lightness perception. *Psychological Review*, *106*, 795-834.
- Gilchrist, A. L. (1980). When does perceived lightness depend on perceived spatial arrangement? *Perception and Psychophysics*, *28*, 527-538.
- Hansen, T., Walter, S., & Gegenfurtner, K. R. (2007). Effects of spatial and temporal context on color categories and color constancy. *Journal of Vision*, *7*.
- Helmholtz, H. (1910). *Helmholtz's Physiological Optics* (Translation from the 3rd German edition ed.). New York: Optical Society of America.
- Helson, H., & Jeffers, V. B. (1940). Fundamental problems in color vision. II. Hue, lightness, and saturation of selective samples in chromatic illumination. *Journal of Experimental Psychology*, *26*, 1-27.
- Hillis, J. M., & Brainard, D. H. (2007). Distinct mechanisms mediate visual detection and identification. *Current Biology*, *17*, 1714-1719.
- Hochberg, J. E., & Beck, J. (1954). Apparent spatial arrangement and perceived brightness. *Journal of Experimental Psychology*, *47*, 263-266.

- Hofer, H., Singer, B., & Williams, D. R. (2005). Different sensations from cones with the same photopigment. *Journal of Vision*, 5, 444-454.
- Jaaskelainen, T., Parkkinen, J., & Toyooka, S. (1990). A vector-subspace model for color representation. *Journal of the Optical Society of America A*, 7, 725-730.
- Kaiser, P. K., & Boynton, R. M. (1996). *Human Color Vision* (2nd ed.). Washington, D.C.: Optical Society of America.
- Katz, D. (1935). *The World of Colour* (R. B. M. a. C. W. Fox, Trans.). London: Kegan, Paul, Trench Truber & Co., Ltd.
- Knill, D., & Richards, W. (Eds.). (1996). *Perception as Bayesian Inference*. Cambridge: Cambridge University Press.
- Kraft, J. M., & Brainard, D. H. (1999). Mechanisms of color constancy under nearly natural viewing. *Proc. Nat. Acad. Sci. USA*, 96, 307-312.
- Krantz, D. (1968). A theory of context effects based on cross-context matching. *Journal of Mathematical Psychology*, 5, 1-48.
- Land, E. H. (1986). Recent advances in retinex theory. *Vision Research*, 26, 7-21.
- Lee, T.-W., Wachtler, T., & Sejnowski, T. J. (2002). Color opponency is an efficient representation of spectral properties of natural scenes. *Vision Research*, 42, 2095-2103.

- Maloney, L. T. (1986). Evaluation of linear models of surface spectral reflectance with small numbers of parameters. *Journal of the Optical Society of America A*, 3, 1673-1683.
- Maloney, L. T. (2002). Statistical decision theory and biological vision. In D. Heyer & R. Mausfeld (Eds.), *Perception and the Physical World* (pp. 145-189). New York: Wiley.
- Maloney, L. T., & Wandell, B. A. (1986). Color constancy: a method for recovering surface spectral reflectances. *Journal of the Optical Society of America A*, 3, 29-33.
- Mamassian, P., & Landy, M. S. (2001). Interaction of visual prior constraints. *Vision Research*, 41, 2653-2668.
- Mamassian, P., Landy, M. S., & Maloney, L. T. (2002). Bayesian modelling of visual perception. In R. P. N. Rao, B. A. Olshausen & M. S. Lewicki (Eds.), *Probabilistic Models of the Brain: Perception and Neural Function* (pp. 13-36). Cambridge, MA: MIT Press.
- Marr, D. (1982). *Vision*. San Francisco: W. H. Freeman.
- McCann, J. J., McKee, S. P., & Taylor, T. H. (1976). Quantitative studies in retinex theory: A comparison between theoretical predictions and observer responses to the 'Color Mondrian' experiments. *Vision Research*, 16, 445-458.

- Motoyoshi, I., Nishida, S., Sharan, L., & Adelson, E. H. (2007). Image statistics and the perception of surface qualities. *Nature*, *Epub ahead of print*; April 18, 2007.
- Newhall, S. M., Nickerson, D., & Judd, D. B. (1943). Final report of the O.S.A. subcommittee on the spacing of Munsell Colors. *Journal of the Optical Society of America*, *33*, 385-412.
- Nickerson, D. (1957). *Spectrophotometric data for a collection of Munsell samples*: U.S. Department of Agriculture.
- Nishida, S., & Shinya, M. (1998). Use of image-based information in judgments of surface-reflectance properties. *Journal of the Optical Society of America A*, *15*, 2951-2965.
- Parraga, C. A., Brelstaff, G., Troscianko, T., & Moorehead, I. R. (1998). Color and luminance information in natural scenes. *Journal of the Optical Society of America A*, *15*, 563-569.
- Parraga, C. A., Troscianko, T., & Tolhurst, D. J. (2002). Spatiochromatic properties of natural images and human vision. *Current Biology*, *12*, 483-487.
- Pratt, W. K. (1978). *Digital Image Processing*. New York: John Wiley & Sons.
- Rao, R. P. N., Olshausen, B. A., & Lewicki, M. S. (Eds.). (2002). *Probabilistic Models of the Brain: Perception and Neural Function*. Cambridge, MA: MIT Press.

- Ripamonti, C., Bloj, M., Mitha, K., Greenwald, S., Hauck, R., Maloney, S. I., & Brainard, D. H. (2004). Measurements of the effect of surface slant on perceived lightness. *Journal of Vision*, 4, 747-763.
- Ruderman, D. L., Cronin, T. W., & Chiao, C. C. (1998). Statistics of cone responses to natural images: implications for visual coding. *Journal of the Optical Society of America A*, 15, 2036-2045.
- Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. *Science*, 237, 1317-1323.
- Shepard, R. N. (1992). The perceptual organization of colors: an adaptation to regularities of the terrestrial world? In J. H. Barkow, L. Cosmides & J. Tooby (Eds.), *The Adapted Mind : Evolutionary Psychology and the Generation of Culture* (pp. 495-532). New York: Oxford University Press.
- Speigle, J. M., & Brainard, D. H. (1999). Predicting color from gray: the relationship between achromatic adjustment and asymmetric matching. *Journal of the Optical Society of America A*, 16, 2370-2376.
- Stiles, W. S. (1967). Mechanism concepts in colour theory. *Journal of the Colour Group*, 11, 106-123.
- Stocker, A. A., & Simoncelli, E. P. (2006). Noise characteristics and prior expectations in human visual speed perception. *Nature Neuroscience*, 9, 578-585.

- van Hateren, J. H. (1993). Spatiotemporal contrast sensitivity of early vision. *Vision Research*, 33, 257-267.
- von der Twer, T., & MacLeod, D. I. A. (2001). Optimal nonlinear codes for the perception of natural colors. *Network: Computation in Neural Systems*, 12, 395-407.
- von Kries, J. (1902). Chromatic adaptation. In D. L. MacAdam (Ed.), *Sources of Color Vision* (pp. 109-119). 1970, Cambridge, MA: MIT Press.
- Wachtler, T., Doi, E., Lee, T.-W., & Sejnowski, T. J. (2007). Cone selectivity derived from the responses of the retinal cone mosaic to natural scenes. *Journal of Vision*, 7, 1-14.
- Wallach, H. (1948). Brightness constancy and the nature of achromatic colors. *Journal of Experimental Psychology*, 38, 310-324.
- Wandell, B. A. (1995). *Foundations of Vision*. Sunderland, MA: Sinauer.
- Weiss, Y., Simoncelli, E. P., & Adelson, E. H. (2002). Motion illusions as optimal percepts. *Nature Neuroscience*, 5, 598-604.
- Wyszecki, G., & Stiles, W. S. (1982). *Color Science - Concepts and Methods, Quantitative Data and Formulae* (2nd ed.). New York: John Wiley & Sons.
- Xiao, B., & Brainard, D. H. (2008). Surface gloss and color perception of 3D objects. *Visual Neuroscience*, in press.

Yang, J. N., & Maloney, L. T. (2001). Illuminant cues in surface color perception: tests of three candidate cues. *Vision Research*, *41*, 2581-2600.

Figure Captions, Bayesian Approaches to Color Vision, David H. Brainard

Figure 1. Information loss in color vision.

A) Color vision begins when light reflects from an object to the eye and is imaged on the retina. The color signal reaching each location of the retina is characterized by its spectrum, which is the amount of power at each wavelength.

B) The spectral sensitivities of the L, M, and S cones. The isomerization rate of photopigment is determined by multiplying the incident power of the color signal at each wavelength by the photopigment's spectral sensitivity and summing over wavelength.

C) Two physically distinct spectra that produce the same isomerization rates in the human L, M, and S cones simultaneously.

Figure 2. Simple example of Bayesian analysis.

A) Likelihood function for simple example. The likelihood function $p(r | i, s)$ is plotted as a function of (i, s) for the case $r = 1$. Values of i and s are assumed constrained between 0 and 4.

B) Prior for simple example. The prior probability distribution $p(i, s)$ is plotted, showing a prior where illuminant values are more likely near $i = 2$ but that is silent about surface reflectance s .

C) The posterior is given by the pointwise product of the likelihood and the prior, and then normalized. The resulting $p(i, s | r)$ is shown for the likelihood in A and prior in B. This example closely follows one provided by Brainard et al. (2008, see their Figure 5)

Figure 3. Prior over surface reflectance

A) Linear model for surfaces. Three basis functions of a linear model for surfaces. The basis functions were obtained through analysis of the principle components of a collection of 462 measured surface reflectance functions (Newhall, Nickerson, & Judd, 1943; Nickerson, 1957).

B) A measured surface reflectance function (solid line) from the same data set used to generate the linear model, and its approximation (dashed line) within the model shown in A.

C) Distribution of model weights for the 462 surface reflectance functions in the data set used to generate the linear model. Each panel shows the histogram for one basis function. The solid lines are a normal approximation. After Brainard and Freeman (1997, see their Figure 4).

Figure 4. Color constancy performance

A) Images of four of seventeen simulated scenes used to compare human performance and a model derived from a Bayesian illuminant estimation algorithm. Each scene has the same spatial structure. In the first three images shown from left to right, the illuminant varies. In the rightmost image, the illuminant is the same as in the second image but the background surface has been changed so that the light reflected from it matches that reflected from the background surface in the leftmost image. Reproduced from Figure 1 of Brainard et al. (2006).

B) Illuminants, achromatic loci, and model predictions plotted in the CIE $u'v'$ chromaticity diagram. This is a standard color representation that preserves information about the relative responses of the L, M, and S cones but not about

intensity. Large open circles show the scene illuminants, with the color key as indicated beneath the images in A. Large filled circles show the achromatic loci measured by observers who adjusted a test patch at the location indicated by the black rectangle in each image. The small open circles show the model's predictions of the achromatic loci. Reproduced from Figure 7 of Brainard et al. (2006).

Figure 5. Trichromatic reconstruction

A) Small patch of sinusoidal isochromatic grating.

B) The intensity of each colored spot represents the isomerization rate of a cone. The class of each cone is indicated by whether it is plotted in red (L), green (M), or blue (S).

C) The isochromatic grating as reconstructed by the Bayesian algorithm. The grating shown corresponds to a spatial frequency of 6 cpd presented at about 1° eccentricity for a human observer; the mosaic is of observer AP of Hofer et al. (2005). Brainard et al. (2008) provide additional reconstruction examples that show similarly veridical performance for low spatial frequency isoluminant gratings and for an additional mosaic.

Figure 6. Small spot experiment

A) Schematic of 5 individual observer cone mosaics. Red, green, and blue circles show locations of L, M, and S cones. Mosaics represent approximately 12' by 12' of visual angle at 1° eccentricity. Reproduced from Figure 2 (top panel) of Brainard et al. (2008).

B) Data from Hofer et al. (2005) for 550 nm spots. Observers named each small spot that they saw and judged namable. The available names were red, orange, yellow, yellow-green, green, blue-green, purple, and white. For each observer, the histogram shows the proportion of each color named used, with the color code corresponding to the name. Note that the white region of the bars represents the proportion of white responses. Not all observers used all available names. Reproduced from Figure 11 of Brainard et al. (2008).

C) Predictions from the Bayesian model, obtained as described in the text, for the experimental conditions corresponding to the data in B. Reproduced from Figure 11 of Brainard et al. (2008). That figure also shows data and predictions for 500 nm and 600 nm spots.

Figure 7. Small spot intuitions

A) A mosaic consisting only of L cones. The white spot in the center indicates a single cone whose stimulation was simulated.

B) This mosaic is identical to the one shown in A, with the exception that the cones surrounding the central L cone have been changed to M cones.

C) Model output when the central cone in A) is stimulated. The result is a bluish-white spot. As described by Brainard et al. (2008), a windowing procedure was applied to model output here and in D) to reduce visible ringing in the reconstruction.

D) Model output when the central cone in B) is stimulated. The result is a redish spot. This figure (A-D) is reproduced from Figure 7 of Brainard et al. (2008).

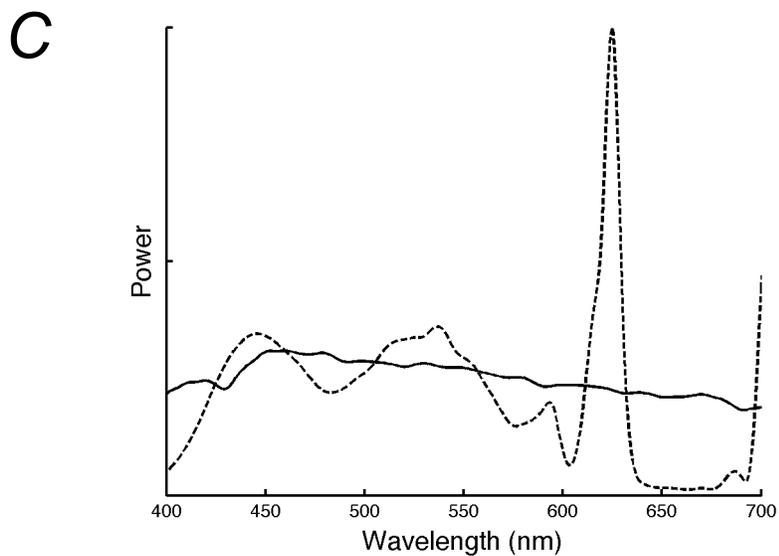
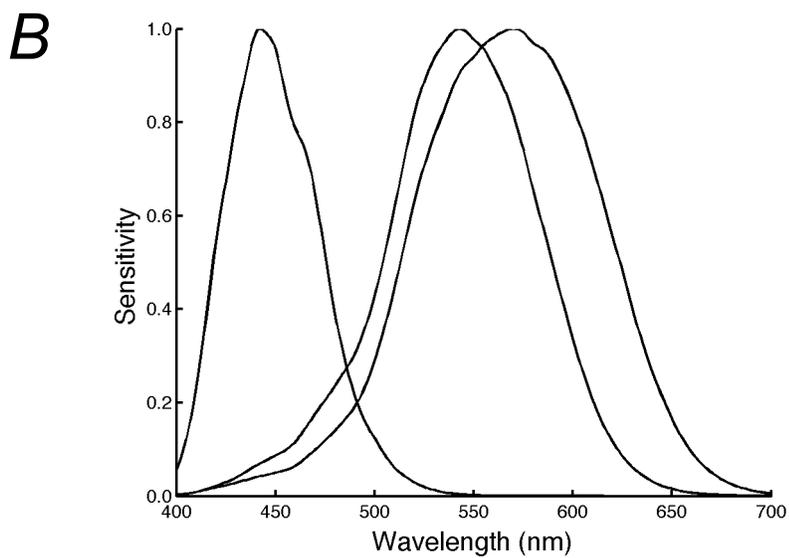
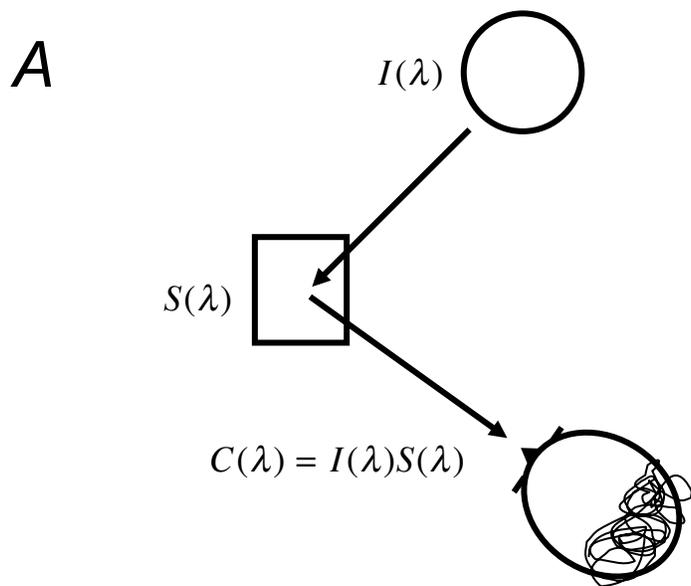
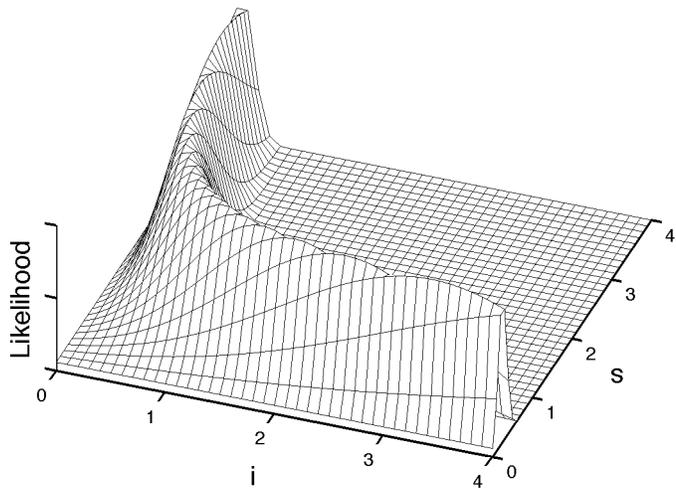
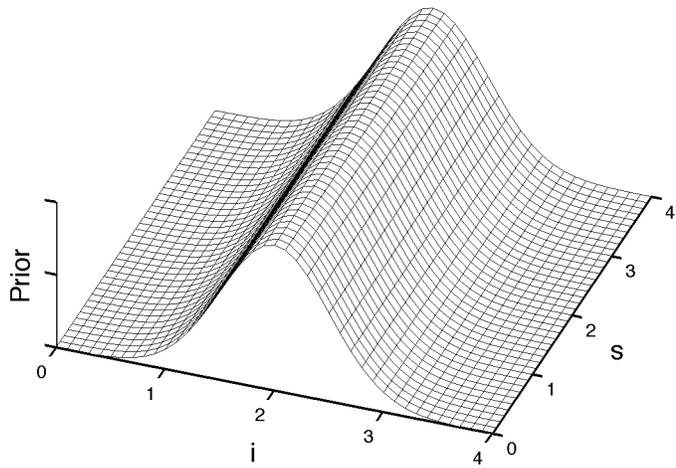


Figure 1

A



B



C

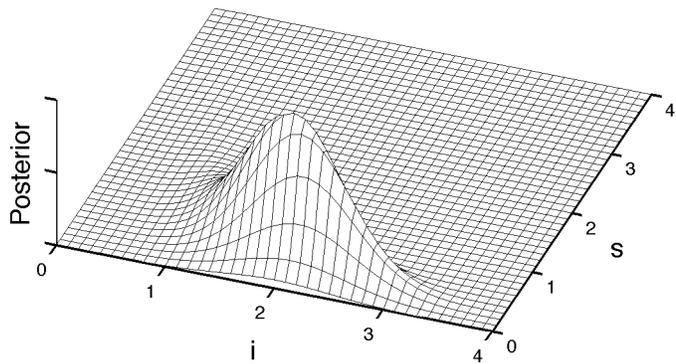
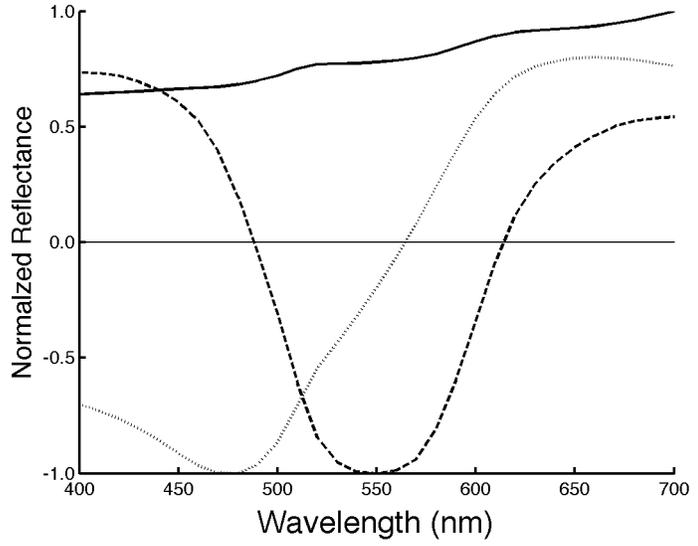
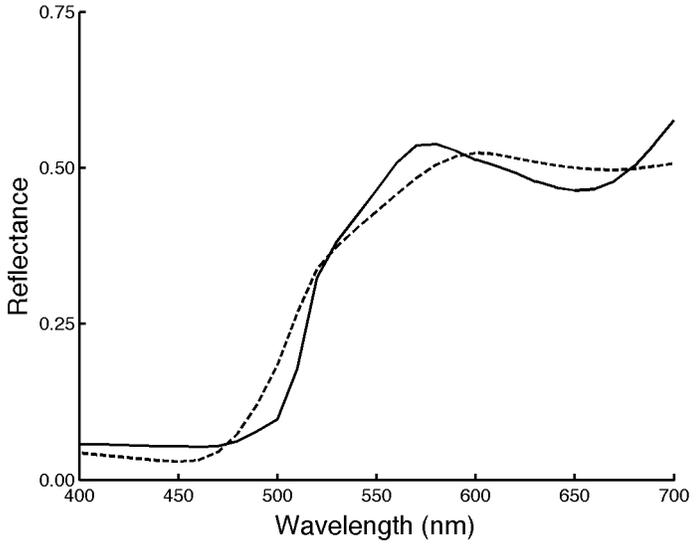


Figure 2

A



B



C

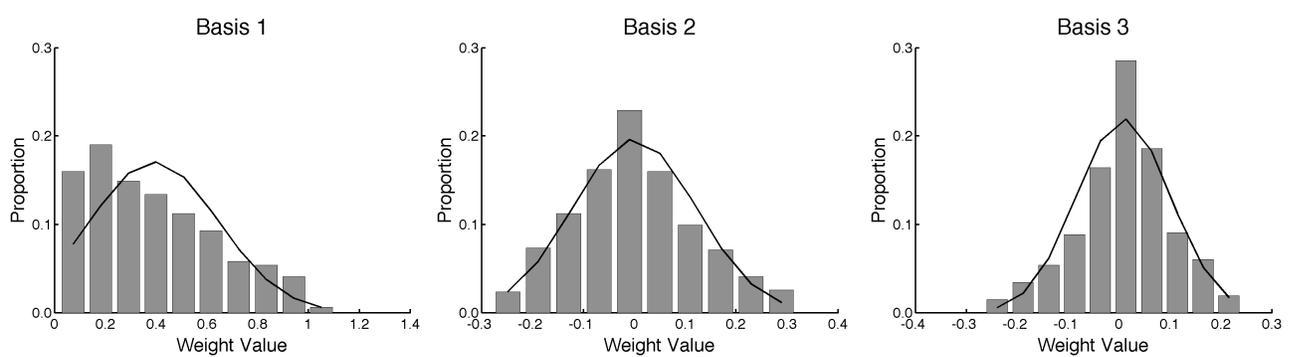
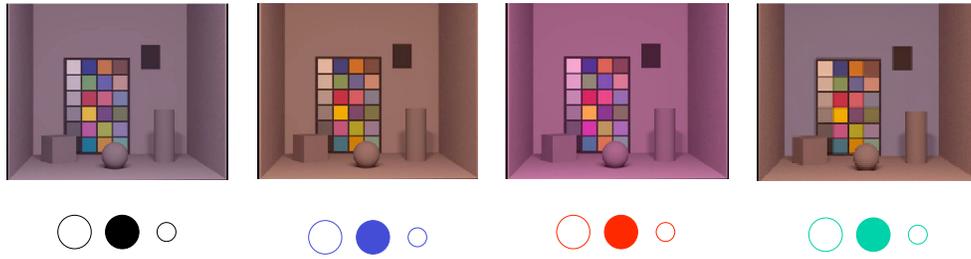


Figure 3

A



B

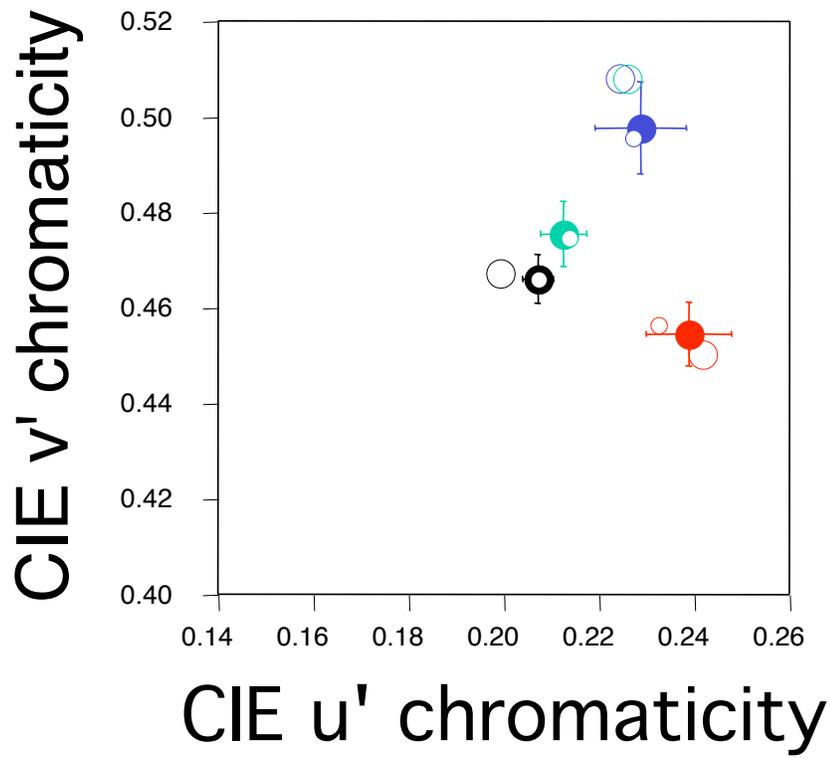
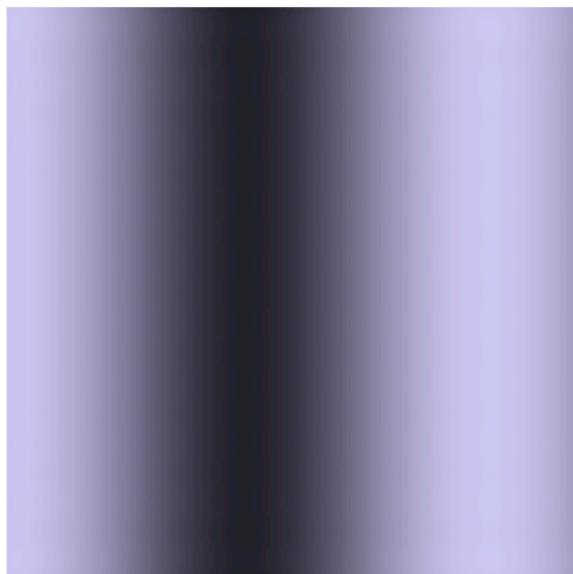
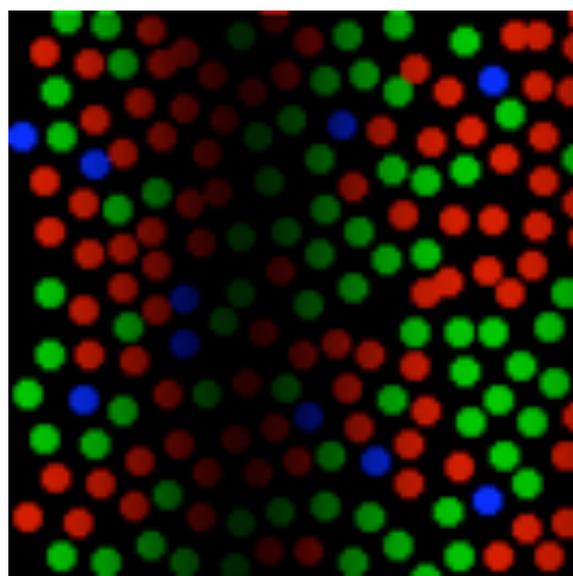


Figure 4

A



B



C

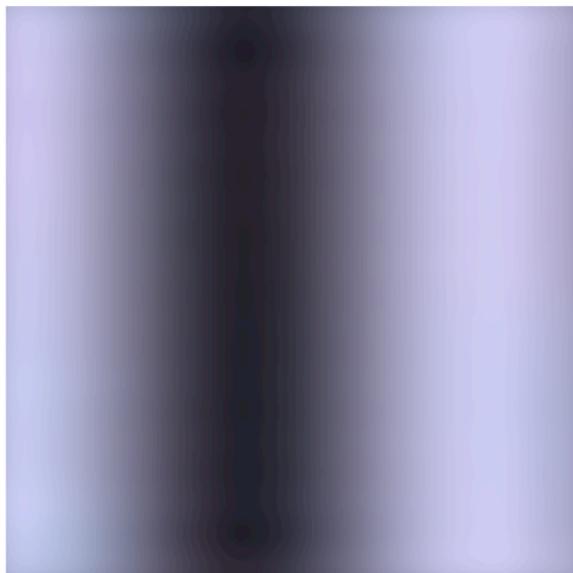


Figure 5

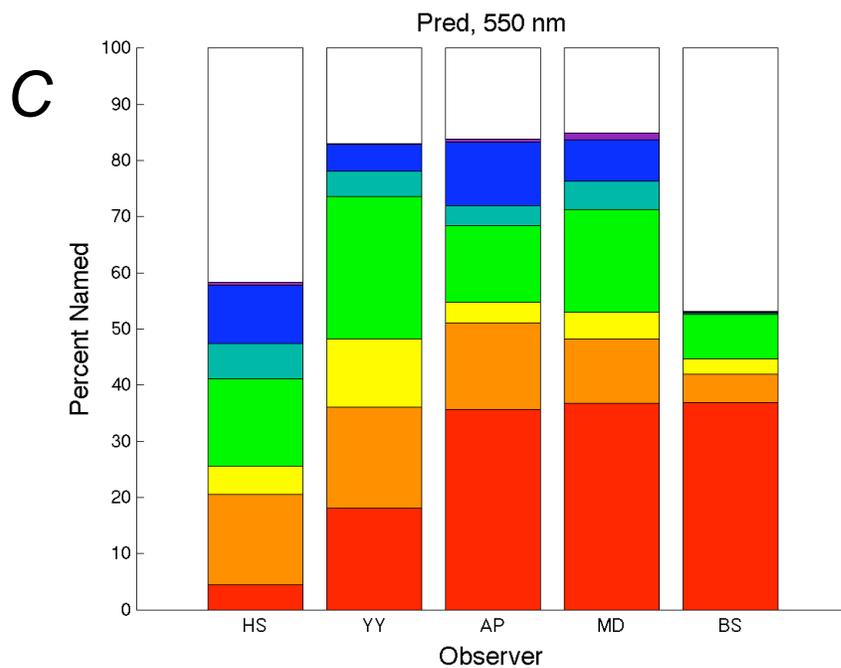
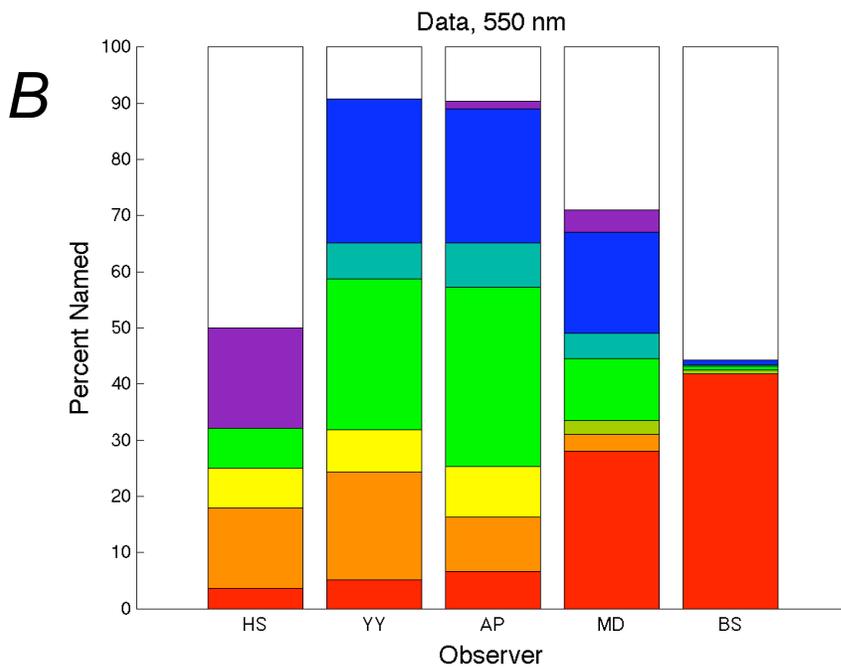
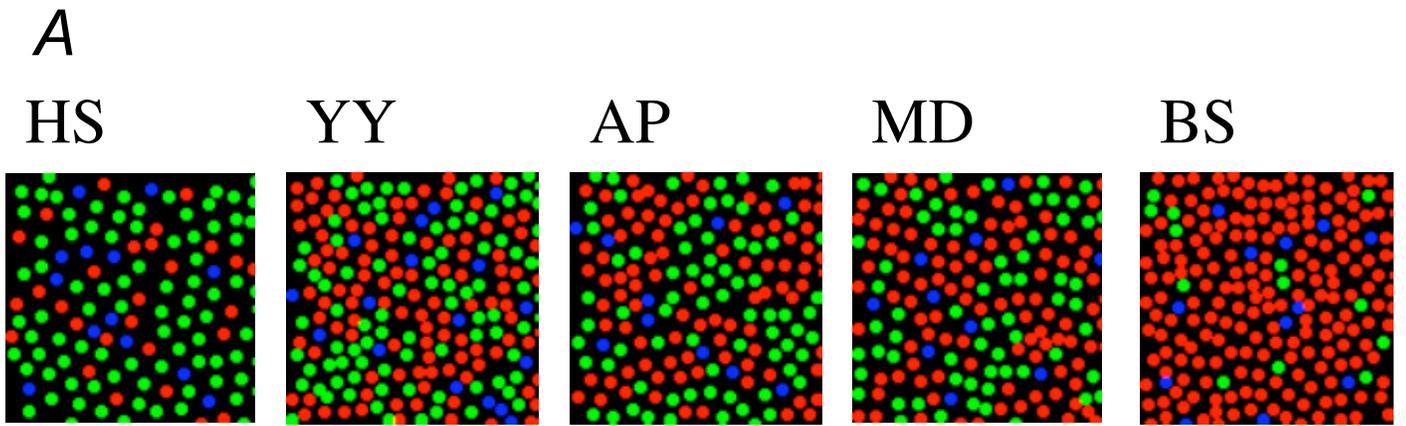
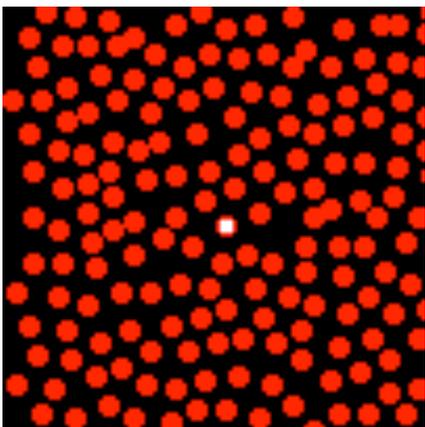
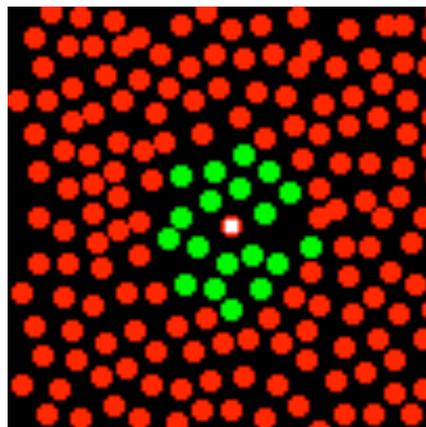


Figure 6

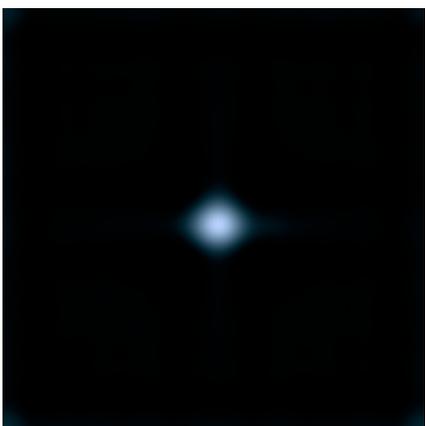
A



B



C



D

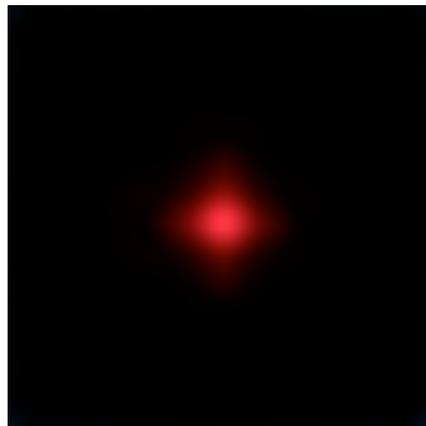


Figure 7