CODING MECHANISMS & CONTEXTUAL INTERACTIONS IN COLOR VISION

CHRISTIAN J. KELLNER

Graduate School of
Systemic Neurosciences
LMU Munich

Dissertation der Graduate School of Systemic Neurosciences der Ludwig-Maximilians-Universität München

5th July 2016

Christian J. Kellner: *Coding Mechanisms & Contextual Interactions in Color Vision*

Reviewer 1: PD Dr. Thomas Wachtler
Reviewer 2: Prof. Dr.-Ing. Stefan Glasauer
Date of oral defense: 8.11.2016

To friends, family and the little loved one

"Life can be magnificent and overwhelming — that is its whole tragedy.
Without beauty, love, or danger it would almost be easy to live."

— Albert Camus

CONTENTS

1

INTRODUCTION

For the Rays to speak properly are not coloured.

— Isaac Newton

The question of what color *is* fascinates philosophers and scientists alike. For philosophers it is of particular interest because it is a phenomenon that everyone has an immediately subjective experience of. Now the question is, do we see things as colorful (the subjectivist view) or are the objects themselves colorful (the realist view)? This is therefore a question of metaphysics, and the answer to what it *is* implicitly answers the question of *where* it is: In the outside world or in our mind and brain. The beginning of the study of color dates back to ancient Greece and natural philosophy. In his work *On Colors* Aristotle postulates that all colors are mixtures of black and white, and identifies "simple colors" with the elements fire, air, water and earth. This theory was influential for a very long time. Modern color science began with Newton's famous refraction studies with prisms, published in 1672. He described his discovery that white light is a composition of colors that can be split up into its constituent parts, its *spectrum*, and then later recombined to again result into white light (Newton, 1671). Newton also marked the turn towards the subjectivist view of color, a position that most color scientists adopted thereafter. According to him, objects have a disposition to reflect light and some exhibit higher refrangibilities for parts of the spectrum than others, a view that corresponds to the modern concept of *reflectance*. These reflected rays "are not [themselves] coloured" but have the "power and disposition to stir up a sensation of this or that Colour" (Newton, 1730). Two more theories are particularly important for our current understanding of color. The first is the theory of trichromacy initiated by Young (1802). Reasoning that not every point in the retina could be sensitive to all sorts of wavelengths ("undulations"), he argued that there should be three sensitive "particles" that would respond to the three primary colors. The ratio of their relative activation ("motion") would then give rise to all possible color sensations. This theory was developed further by von Helmholtz (1867), who postulated the existence of three "nerve fibers" with broad and

overlapping spectral sensitivities, with each fiber maximally sensitive to a specific wavelength. The other important theory underlying our current understanding of color is the theory brought forward by Ewald Hering in 1878 (Mollon, 2003). He started his considerations about color with the subjective experience that it is impossible to perceive a yellowish blue or a greenish red. He therefore concluded that three opponent processes must exist somewhere in the visual system, consisting of blue vs yellow, red vs green and black vs white.

Nowadays we know that both theories — that at the time seemed to be mutually exclusive — are at least partially right and can be combined. With the subsequent discoveries of the photoreceptors and the opponent processes (De Valois, Smith, Kitai, & Karoly, 1958; De Valois, 1960; Svaetichin & Macnichol, 1959), the current view is that color vision is enabled by three different cone photoreceptors in the retina, whose signal is then processed in an opponent manner. From there, the signals are transmitted to the primary visual cortex of the brain for further processing. These findings also established *color vision* as a discipline of neuroscience, the study of the brain and the nervous system.

In the light of neuroscience, color is information that can be used by the brain for adequate action and adaptive behavior, particularly in an ever changing environment. Indeed, we heavily rely on vision as indicated by the size of the visual cortex — in humans more than a third of the brain is devoted to the processing of visual information. How the nervous system processes information is therefore also an important question for *color vision*. The systemic study of the brain "in terms of the information processing properties of structures that make up the nervous system" (Churchland, Koch, & Sejnowski, 1988) is the central goal of *computational neuroscience*. It uses theoretical analysis, mathematical models and computational simulations to shed light on how individual neurons, neuronal circuits, and ultimately the whole brain encode and process information. Since information processing is related to the activity of neurons, there is a constraint to only encode useful information. Therefore, signals transmitted by the sensory organs to the rest of the nervous system should be encoded efficiently to not waste energy. This introduces the notion that organisms and their nervous system should be adapted to the environment they are living in.

So far, the study of *color vision* in neuroscience has largely focused on information processing at the level of the retina. In this work we use computational simulations together with psychophysical experiments to gain new insights into the processing of color in the early visual cortex.

But first, the basic concepts of coding mechanisms will be introduced, with an emphasis on efficient coding (Section 1.1); this is followed by a brief overview on how information is processed in subcortical pathways (Section 1.2) — especially in the light of efficient coding — and lastly an overview of what is known about the processing of color in the visual cortex so far (Section 1.3). This will be accompanied by an explanation of some perceptual effects that guided the development of models of cortical processing in this thesis.

1.1 CODING MECHANISMS AND PRINCIPLES

Neurons are specialized cells that can receive, process and send information to each other via electrical and chemical signals. Most communication in the nervous system is realized via action potentials, also called spikes, which are brief stereotypical binary events. Neurons can use these spikes to encode information in a number of ways, most prominently via the spike rate, i.e. the frequency of the spikes, or the spike timing (see for example Gerstner, Kreiter, Markram, and Herz, 1997).

To study *what* neurons encode we have two choices. We can generate models of information processing, make predictions and test whether neurons behave as expected. Alternatively, we can present stimuli to the sensory systems of an organism and record which stimuli evoke what kind of response in the neurons, if any. In now famous experiments, Hubel and Wiesel (1962) recorded from neurons in the visual cortex, later called "simple cells", that responded to specific orientations of bars and gratings. The strength of the neurons' response depended on the orientation and the precise location of the stimulus in the visual field. The region in space and time of the "sensory surface" that a neuron responds to is called its *receptive field* (Hartline, 1938; Sherrington, 1906) and the dependency of the neuron's response on stimulus features is called its *tuning*.

Another important dimension of neural coding is the number of neurons involved in the coding of a variable, i.e. how sparse a neural code is. One extreme is that a single neuron encodes a specific feature, for example the identity of a person. This led to the term "grandmother neuron", the idea that the activity of a single neuron encodes a rather complex object, e.g. your grandmother (Gross, 2002). And indeed, Quiroga, Reddy, Kreiman, Koch, and Fried (2005) have found cells in the human medial temporal lobe (MTL) each of which responded to different pictures and the name of specific landmarks, persons, and other objects. This is

of course no proof that the found cells were really "grandmother cells" because neither were the cells exposed to all possible complex objects known to the subject (so it is not certain the cell would only fire to a single concept), nor could the experimenters record from all cells in MLT at the same time (so it is not certain that other neurons would not respond to the same stimulus). However, this result suggests a very sparse and explicit encoding of complex features.

1.1.1 *Population Coding*

A coding scheme that involves more than one neuron to encode a single variable is often called an ensemble code, or population code. Its basic building blocks are neurons that are all tuned to a certain feature with broad and overlapping tuning curves, and prefer one value of the feature to which they respond to maximally. The "simple cells" found by Hubel and Wiesel (1962) constitute such a population code, where orientation is encoded by a population of simple cells. Each of these cells has a preferred direction it responds maximally to, and its response declines with the angular distance of the presented bar to the preferred orientation.

The standard coding model for neurons — often called *units* in abstract, idealistic mathematical models — with cosine tuning curves is the *population vector* (Georgopoulos, Schwartz, & Kettner, 1986). A certain value x, expressed in vector coordinates of the feature space, is encoded by the vector sum of the preferred direction (p_i) of each unit i that is part of the population code, weighted with its response f_i:

$$\vec{x} \propto \sum_i f_i \vec{p_i} \tag{1.1}$$

In general, population coding schemes are particularly robust, as the removal of a few cells does not result in a total coding failure (Pouget, Dayan, & Zemel, 2000). Population codes are used by the brain in many places beyond the visual system, for example in the control of saccadic eye movements (C. Lee, Rohrer, & Sparks, 1988), arm movements (Georgopoulos et al., 1986), the auditory system (Grothe, 2003) and the representation of space (Giocomo, Moser, & Moser, 2011).

1.1.2 *Efficient Coding*

A theoretical framework to study neural coding that makes specific predictions of how neurons should code information is the theory of *efficient coding*. The basic assumptions are that organisms should be adapted to the environment they live in, and that information processing by the nervous system comes with a cost and should therefore be efficient. This specifically applies to the organism's sensory systems, which should be adapted to the statistics of their respective inputs (Barlow, 1961; Attneave, 1954).

Information Theory

The mathematical basis of the efficient coding hypothesis is information theory, which is the quantitative study of information. Since its development by Shannon (1948), it has had a profound impact on many scientific fields. Information theory is based on probability theory and statistics. The basic building blocks are messages that consist of symbols from an alphabet. Messages are sent from a source through a channel with possible noise to a recipient. The *information content* of a message m, chosen from a finite ensemble of messages M, is defined as (Shannon, 1948; MacKay, 2003):

$$h(m) = \log_2 \frac{1}{p(m)} = -\log_2 p(m) \tag{1.2}$$

Intuitively this is a measure of the "surprise", or "unexpectedness" of the received message. It is measured in bits[1]. The information content is independent of the semantic content of the message, i.e. what receiving the message means for the recipient. The central concept of information theory is that of the *entropy* of a message ensemble M, which is the finite space of all possible messages. It is defined as:

$$H(M) = -\sum_{m \in M} p(m) \log_2 p(m) \tag{1.3}$$

It represents the uncertainty that a specific message will be received and is a measure of the average information content per message. There is a direct link between entropy and efficiency: information transmission

1 The unit bits comes from the choice to use the $\log_2$. Other choices are possible, such as the natural logarithmic, then measured in nats.

will be most efficient when entropy is maximal, which means that all messages are equally likely to occur, or in other words, the probability distribution of M is uniform. If, and only if, the same message is chosen ($p(m) = 1, m \in M$) is the entropy ($H(m)$) zero. The *conditional entropy* of A and B is the average uncertainty we have of A given we know B:

$$H(A|B) = H(B, A) - H(B) \tag{1.4}$$

where, $H(B, A)$ is the *joint entropy* of A and B. A direct measure of the "redundancy" between A and B is the *mutual information* between the two, defined as:

$$I(A; B) = H(A) - H(A|B) \tag{1.5}$$

It measures the reduction in the uncertainty about A we get on average if we know B, or conversely the amount of information that A tells us about B. If A and B are independent then $H(A|B) = H(A)$ which means $I(A; B)$ is zero for that case.

Single Neurons

What does efficient coding mean for a single neuron? A typical neuron receives input from a number of synapses, integrates these inputs and then generates a single output in terms of action potentials. This imposes a natural constraint on the distribution of outputs, since there is a biologically imposed upper bound on the number of action potentials a neuron can sustain per second (the firing rate of neurons).[2] In general, information transmission is efficient when the entropy output is maximal (Shannon, 1948; Laughlin & Hardie, 1978; Laughlin, 1981; Nadal & Parga, 1994). In the case of a single neuron with limited firing rates, this means that the distribution of firing rates should be uniform, i.e. all firing rates should be equally likely. Figure 1.1 shows an example of a simple case with only one input i that has a probability density of $p(i)$ and a neuron with a sigmoidal transfer function $t(i)$. Optimal coding can be achieved when the slope of the transfer function $t(i)$ is proportional to the probability density $p(i)$ of the input. In this case the output distribution $p(o)$ is uniform.

2 Various other coding schemes exist in the nervous system. Photoreceptors, for example, release a neurotransmitter and bipolar cells in the retina communicate via graded potentials.

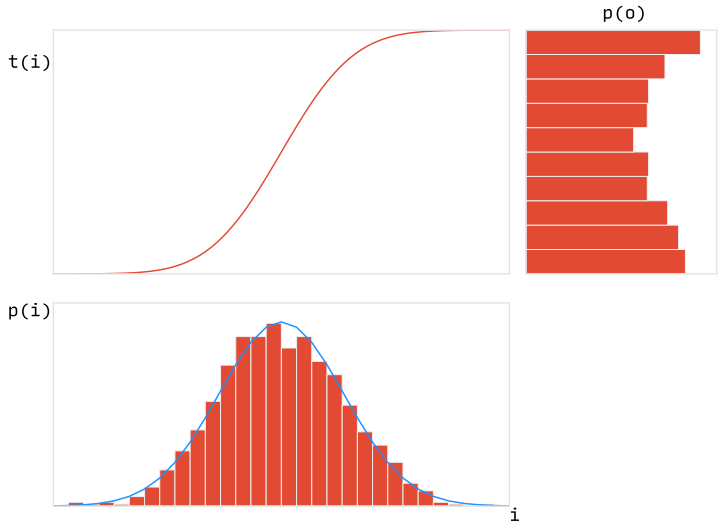


Figure 1.1: Optimal transmission of information in a single neuron. For a given input i the information transfer in the neuron is efficient when the entropy of the output o is maximal, i.e. all outputs are equally likely and the output distribution $p(o)$ is flat. This is achieved when the slope of the transfer function $t(i)$ is proportional to the probability density of the input $p(i)$.

Multiple Neurons

In the case of multiple neurons the reasoning of the single neuron case still applies, but additionally requires the joint output to be taken into account. There is redundancy in the overall signal if information encoded by one neuron is also encoded by others. In other words, knowing the response of one neuron should not tell us anything about the response of another neuron, i.e. they should be statistically independent.

1.1.3 *Finding Efficient Codes*

If we assume a generative model where a signal $\mathbf{x}$ is the result of the linear combination of source signals $\mathbf{s}$, i.e. $x_i = \sum_j a_{ij} s_j$, or in matrix notation,

$$\mathbf{x} = \mathbf{A}\mathbf{s} \tag{1.6}$$

then the task of finding an efficient code is to decompose the signal $\mathbf{s}$ into a neural code $\mathbf{o}$ where there is as little redundancy as possible in the code. In mathematical terms, assuming a simple linear network:

$$\mathbf{o} = \boldsymbol{W}\boldsymbol{x} \tag{1.7}$$

Decorrelation Methods

In theory, redundancy reduction can involve statistical dependencies of any order. If we restrict the reduction of redundancies to second-order dependencies, we *decorrelate* or *whiten* the signal. The amount of correlation in the signal can be approximated by the covariance matrix $\boldsymbol{\Sigma}$. In terms of the output signal, assuming the signal has zero mean:

$$\boldsymbol{\Sigma}_{\mathbf{o}} = \langle \mathbf{o}\mathbf{o}^{\intercal} \rangle \tag{1.8}$$

Decorrelation has been achieved if covariance between signals vanishes. This means that the covariance matrix becomes diagonal:

$$\boldsymbol{\Sigma} = \text{diag} \tag{1.9}$$

For our case (equation 1.7), with $\boldsymbol{\Sigma}_{\mathbf{o}} = \mathbf{I}$ (Bell & Sejnowski, 1997):

$$\boldsymbol{W}^{\intercal}\boldsymbol{W} = \langle \boldsymbol{x}\boldsymbol{x}^{\intercal} \rangle^{-1} \tag{1.10}$$

However, this does not impose a unique solution on the matrix $\boldsymbol{W}$. Some special solutions do exist, however, such as Principal Component Analysis and Zero-phase Component Analysis.

Principal Component Analysis, the orthogonal solution, is obtained via eigenvalue decomposition of the covariance matrix, $\langle \boldsymbol{x}\boldsymbol{x}^{\intercal} \rangle = \mathbf{E}\boldsymbol{\Lambda}\mathbf{E}^{-1}$, where the rows of $\mathbf{E}$ are the eigenvectors and the entries on the diagonal of $\boldsymbol{\Lambda}$ the eigenvalues. From this we obtain $\boldsymbol{W}_{\mathrm{PCA}} = \boldsymbol{\Lambda}^{-\frac{1}{2}}\mathbf{E}^{\intercal}$. This corresponds to a rotation of the data, so that the new axes in the coordinate system — the principal components — capture the direction of most variance in the data. The filters of PCA, the rows of $\boldsymbol{W}_{\mathbf{PCA}}$, are global filters (Bell & Sejnowski, 1997).

Zero-phase component analysis is the symmetrical solution, i.e. $W = W^{\intercal}$. The whitening matrix is then: $\boldsymbol{W}_{\mathrm{ZCA}} = \langle \boldsymbol{x}\boldsymbol{x}^{\intercal} \rangle^{-1/2}$ (Bell & Sejnowski, 1997). In contrast to PCA, the filters of ZCA are local and, interestingly, have a center-surround like structure similar to the receptive fields of retinal ganglion cells (cf. Section 1.2). Data filtered with ZCA is as close as possible to the input data (Eichhorn, Sinz, & Bethge, 2009).

Independent Component Analysis

Another method to reduce redundancies, which also takes higher order statistics into account, is Independent Component Analysis (ICA). It was developed as a solution to the blind source separation problem. Applied to our generative model, the goal is to recover the sources $\mathbf{s}$ of the mixed signal $\mathbf{x}$, but neither the sources nor the mixing of the sources $\boldsymbol{A}$ are known. Remarkably, we can still approximate a solution if the sources $\mathbf{s}$ are statistically independent and non-Gaussian. ICA will try to recover the sources via a linear filter matrix $\boldsymbol{W}$ so that the recovered data $\mathbf{o}$ is as statistically independent as possible. It is therefore not constrained to second-order statistics but takes higher order statistics into account. Since its goal is to find as statistically independent signals as possible, ICA is an excellent tool to find efficient codes because statistical independence is precisely the desired property of outputs in the case of multiple coding neurons (see Section 1.1.2 above). Formulations of ICA in the context of efficient coding were developed in terms of artificial neuronal networks based on information maximization (*infomax* algorithm, Linsker, 1988; Bell and Sejnowski, 1995): We want to learn a linear transformation $\boldsymbol{W}_{\mathbf{ICA}}$ that maximizes the joint entropy of the output after passing it through a nonlinear transfer function $H(t(\mathbf{o}))$. Following the entropy gradient the parameters of the nonlinearity will be adjusted. A corresponding learning rule for $\boldsymbol{A}$ is presented in Chapter 2. The columns of $\boldsymbol{A}$, with $\boldsymbol{A} = \boldsymbol{W}_{\mathrm{ICA}}^{-1}$, are called the basis functions and can be compared to the receptive fields of a neuron.

1.2 SUBCORTICAL PROCESSING OF COLOR

The processing of color information by the visual system begins in the retina, where light is converted into neural signals via special sensory cells called photoreceptors. Humans have two classes of photoreceptors, the rods and the cones. Rods have a higher sensitivity and are mostly active under low light conditions (scotopic vision). Cones on the other hand are less sensitive; while they have a poor signal in low lights, they are active during normal daylight (photopic vision). The sensitivity of the photoreceptor is based on the expression of a photosensitive pigment, the opsin, which in cones is located in the outer segment. The specific spectral sensitivity of the opsin is determined by its amino acid composition, which is genetically regulated. Most mammals possess two different types of cones with different sensitivities: one cone that is sensitive to

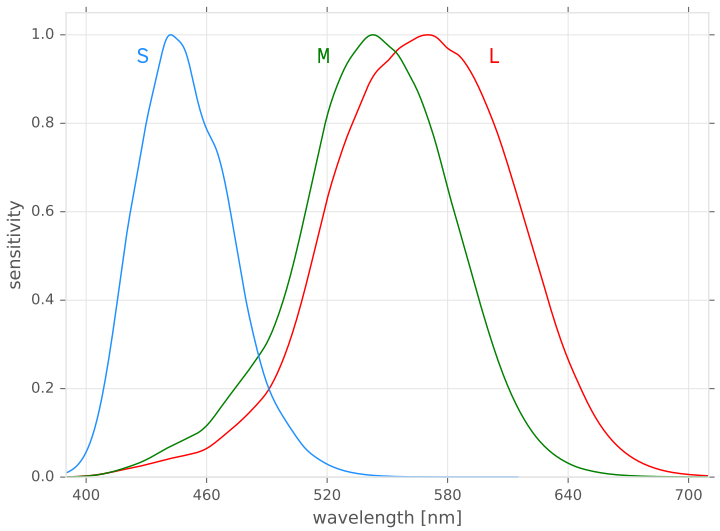


Figure 1.2: Sensitivities of S- (blue), M- (green), and L- (red) cone photoreceptors of the human retina. Data from Stockman and Sharpe, 2000.

the short wavelength part of the spectrum and another that is sensitive to long wavelengths. This forms the basis of color vision in dichromatic animals. The gene of the long wavelength sensitive opsin is located on the X chromosome. In humans and other primates — old world monkeys and apes — the duplication and mutation of this gene led to a third opsin type. They differ only by a few amino acids, which corresponds to a difference of around 30nm in their spectral sensitivities. In total, we therefore have three different types of cones: The S-cone, which is most sensitive to short wavelengths with a peak around 430nm , and two cones most sensitive to longer wavelengths, the M- and L-cones, with peaks around 530nm and 560nm , respectively. All cones have a broad sensitivity range of around 100nm . The spectral sensitivity curves are depicted in Figure 1.2. The presence of these three different cone types forms the basis of our trichromatic color vision.

The spatial layout of cones in the retina — the so-called cone mosaic — is non-uniform with a sharp peak in the density of cones in an area called the *fovea*. Comparisons of cone densities across species with different eye sizes revealed similar densities, suggesting that cone packing in the fovea has reached its limit. The gain of trichromatic color vision was a trade-off: the introduction of a third photoreceptor reduced the spatial resolution and decreased the signal-to-noise ratio (Williams, Sekiguchi,

Haake, Brainard, & Packer, 1991). Clearly the benefits of trichromatic color vision must have outweighed the loss in spatial resolution. The distribution of cone types is also non-uniform across the cone mosaic. S-cones are sparse and only amount to around 5-10% of all cones (Curcio et al., 1991). The predominant types are therefore L- and M-cones.

The signal of photoreceptors corresponds to the amount of photons they absorb over a period of time, also called the quantum catch. The spectral sensitivity determines the probability to absorb a photon of a specific wavelength. After the absorption has taken place, the information on the photon's wavelength is lost. As a result, it is impossible to tell changes in wavelength and intensity apart if only the signal of one photoreceptor type is taken into account. This is called the *principle of univariance* (Rushton, 1972). To enable color vision — which by definition is the ability to discriminate different spectral compositions of light regardless of their intensity — it is therefore necessary to compare the signals of more than one cone type.

The basic structure of these cone signal comparisons, as discovered by Kuffler (1953), is that of a center-surround antagonism, formed by lateral inhibition (Hartline, 1969): The signals of a center region in the retina are combined with an opposing signal from the surrounding area. Cone opponency starts at the cone-bipolar synapse, with negative feedback from horizontal cells. This in turn determines the receptive fields of the retinal ganglion cells, to which bipolar cells connect. The retinal ganglion cells constitute the output layer and their axons form the optic nerve, which transports signals to the lateral geniculate nucleus (LGN) in the thalamus. There are ON and OFF variants of retinal ganglion cells: in ON cells, light in the center region results in excitation of the cell, while light in the surround results in inhibition; in OFF cells, it is the other way around, i.e. light in the center region inhibits the cell, while surround stimulation excites the cell. There are three different groups of retinal ganglion cells, forming three different parallel pathways. They are named after the region of the LGN they project to.

The magnocellular pathway originates in parasol ganglion cells, which sample randomly, but almost exclusively from L- and M-cones with negligible input from S-cones (S. C. S. Lee & Grünert, 2007). Their receptive field is large and both ON and OFF variants exist. The responses of parasol cells are transient and their sensitivity to achromatic contrast is high (B. B. Lee, 2011). They exhibit no wavelength selectivity.

The parvocellular pathway begins with the midget ganglion cells which, again, exist in ON and OFF variants. Their center-surround areas are small and in the fovea the center receives input from a single cone (Boycott, Dowling, & Kolb, 1969). The exact composition of the surround is still an open question (Lennie, Haake, & Williams, 1991; Reid & Shapley, 1992; B. B. Lee, Kremers, & Yeh, 1998; Martin, Lee, White, Solomon, & Rüttiger, 2001; Reid & Shapley, 2002; Buzás, Blessing, Szmajda, & Martin, 2006; G. D. Field et al., 2010; Crook, Manookin, Packer, & Dacey, 2011; B. B. Lee, Shapley, Hawken, & Sun, 2012) but even random selection of L- and M-cones for the surround would lead to an L-M, or M-L cone opponency (Paulus & Körger-Paulus, 1983). Outside of the retina, their center region is larger but still exhibits a bias towards one cone type. Therefore, non foveal midget ganglion cells are also cone-opponent. They prefer chromatic over achromatic stimulation and encode spatial and chromatic information.

The S-cone signals are encoded by small bistratified ganglion cells (Crook et al., 2009), which receive ON input from S-cones and OFF input from a combination of L- and M-cones. They are therefore S-LM color-opponent. Small bistratified ganglion cells project to the koniocellular layer of the LGN (Tailby, Szmajda, Buzás, Lee, & Martin, 2008). S-cone OFF signals can be found in the LGN (Derrington, Krauskopf, & Lennie, 1984; Valberg, Lee, & Tigwell, 1986; Szmajda, Buzás, FitzGibbon, & Martin, 2006) but their origin in the retina remains unclear. It has been speculated that an intrinsically photosensitive ganglion cell, which expresses a photosensitive pigment, could be a possible candidate (Dacey et al., 2005). See also Dacey, Crook, and Packer (2013) for an in depth review of S-OFF (and S-ON) cone-opponent pathways.[3]

1.2.1 *Cone-opponent Colorspace*

Specifying color requires a colorspace: a metric and organization of color in a coordinate system. There are many different colorspaces that follow different distinct organization principles. The first to establish a quantitative link between the visible light spectrum and human color perception is the CIE 1931 colorspace (Smith & Guild, 1931). Although widely used, this colorspace has certain drawbacks for vision research: it is not perceptually uniform, does not provide a metric for discrimination, and

3 The picture of the pathways presented here provides only a brief overview. B. B. Lee (2014), for example, provides a more in depth description with an emphasis on open issues.

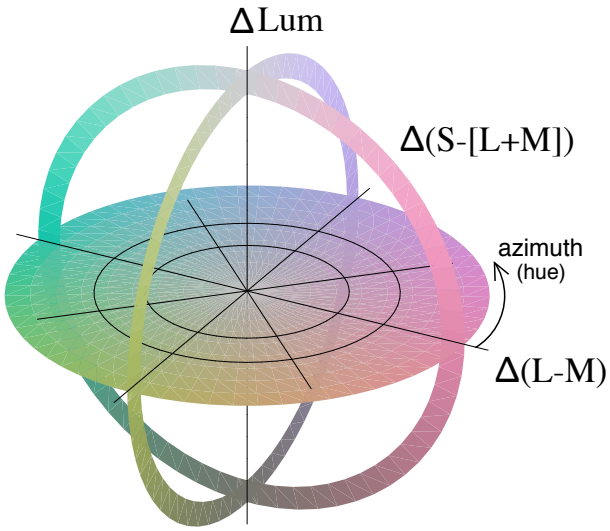


Figure 1.3: Color space formed by the three parallel pathways: axes correspond to variations in L+M, L-M and S-(L+M) signals. The azimuth angle ϕ is a direct measure of the color quality, the hue. Elevation θ corresponds to luminance and saturation varies with the radius, i.e. the distance from the origin.

most importantly, the relation of coordinates in the colorspace to cone excitation is "obscure" (Boynton, 1996). For color vision research it seems more appropriate and helpful to use a metric based on the physiological properties of the visual system. The colorspace developed by MacLeod and Boynton (1979) provides a direct measure of relative cone excitations. Later, Derrington et al. (1984) developed a cone-opponent contrast space based on the retinal pathways that incorporates the fact that neurons encode deviations from the mean rather than an absolute value. A similar colorspace, depicted in Figure 1.3, is used to specify chromaticities in this thesis. The three axes are defined by the three pathways. The vertical axis is the sum of L+M signals and corresponds to the luminance. Orthogonal to that are two color-opponent axes. The first represents the difference between L- and M-cone signals, while the second, again orthogonal to both previous axes, represents the differences between S and the sum of L and M cone signals. All coordinates in the colorspace are deviations from the origin, which is the reference point normally set to a neutral gray (which stimulates all three cones equally). Coordinates are usually specified in polar form: changes in elevation θ correspond to luminance changes. Points in the plane with zero elevation are isoluminant, i.e. they have the same luminance as the reference point. Changes in azimuth ϕ correspond to changes in hue. The distance from the origin in an isolumi-

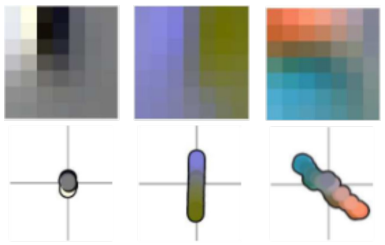


Figure 1.4: Top| Example basis functions. Pixel R, G, B values correspond to relative activations of the L,M, and S cones of each pixel. Bottom| Chromaticities of single pixels for each basis function are plotted in the cone-opponent color space (cf Figure 1.3). Plots from data of T.-W. Lee et al. (2002)

nant plane corresponds to chromatic contrast from the reference point which, for neutral gray, is equal to saturation.

1.2.2 *Efficient Coding in the Retina*

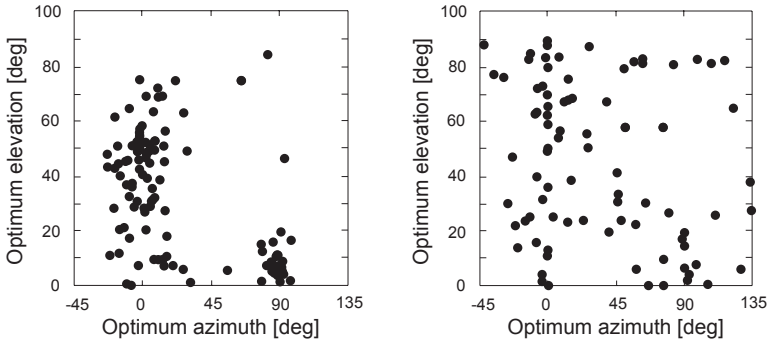
The visual system is confronted with two possible sources of redundancy. Firstly, the sensitivity curves of the cones are broad and overlapping (cf. Figure 1.2), especially for the L- and M-cone systems. Secondly, the image that falls onto the retina originates from physical objects with mostly uniform reflectances, leading to a high correlation of nearby pixel values (Dong & Atick, 1995; D. J. Field, 1987). Additionally, natural scenes are characterized by relatively smooth spectra, which further contribute to high correlations between cone signals (T.-W. Lee, Wachtler, & Sejnowski, 2002; Maloney, 1986).

It has long been argued that the retina should perform a decorrelation of the input signal (Atick, 1992; Doi & Lewicki, 2014), particularly the center-surround structure (Atick & Redlich, 1993, 1990; Srinivasan, Laughlin, & Dubs, 1982). Recently, empirical studies have confirmed that the retina does indeed remove redundancies, but results also highlighted the importance of contributions from the non-linearity stage (Pitkow & Meister, 2012; Tokutake & Freed, 2008; Doi et al., 2012).

How does opponency contribute to efficient coding? Buchsbaum and Gottschalk (1983) performed Principal Component Analysis (PCA, cf. Section 1.1.3) on the spectral sensitivities of the human cones. They obtained one achromatic and two color-opponent "channels" with red-green and

blue-yellow opponency — a striking similarity to the three subcortical parallel pathways. Ruderman, Cronin, and Chiao (1998) built on this work, now by taking natural statistics into account. Hyperspectral images, where the entire visible light spectrum is sampled in small steps (e.g. 2nm) for each pixel, were taken to derive L-, M-, and S-cone activations for natural ("foliage dominated") scenes. PCA was performed on these cone signals. They found three directions in the (logarithmic[4]) cone space that robustly decorrelated the natural scenes. The first one corresponded to luminance changes while the other two corresponded to reddish-greenish and blueish-yellowish opponency. Both studies used PCA, a decorrelation method that results in an orthogonal transformation, and thus opponency was already implied in the algorithm (Buchsbaum & Gottschalk, 1983; Ruderman et al., 1998). Therefore, Wachtler, Lee, and Sejnowski (2001) used Independent Component Analysis (ICA) — a method that does not enforce an orthogonal basis but rather tries to make the outputs as statistically independent as possible (cf. Section 1.1.3) — to learn a sparse and efficient code for L-, M-, S-cone signals derived from hyperspectral images of natural scenes. The resulting basis functions could be categorized into three different types: the first type consisted of achromatic basis functions that were oriented and localized, similar to the simple cell-like edge detectors found in previous studies that used similar methods but on grayscale images (Olshausen & Field, 1996; Bell & Sejnowski, 1997). The second type was homogeneous chromatic and the third was chromatic opponent, i.e. color edge detectors. When the LMS pixel values of the chromatic opponent basis functions were plotted in the cone opponent colorspace, they fell on a line that crossed through the origin. There were two subtypes: one where the axis coincided with the blue-yellow opponency axis of the colorspace, and a second with an oblique axis that possessed orange-teal color opponency. Figure 1.4 shows three example basis functions, one achromatic and one of each of the cone-opponent subtypes. A follow-up analysis with the same method and dataset but with non-overlapping cone spectra by T.-W. Lee et al. (2002) produced similar basis functions that were also cone-opponent, except that the group of basis functions that showed orange-teal opponency was less tilted and closer to the L-M axis. The findings of both studies strongly suggest that cone-opponency results from the statistical structure of natural scenes.

4 Actual signals used in the study were logarithmically transformed before the PCA step; this was done to improve the processing by PCA, but additionally justified by the Weber-Fechner law.



(a) Lateral geniculate nucleus.

(b) Striate cortex.

Figure 1.5: Color preferences of neurons in the LGN (a) and striate cortex (b) measured in the macaque monkey. Both panels show the distribution of stimulus parameters, hue (azimuth) and luminance (elevation), that elicited the maximum response. In the LGN, preferred hues are clustered around the two cardinal axes of the colorspace, corresponding to the two chromatic pathways. In striate cortex, the distribution of hue preferences is more widely distributed. Data from Lennie et al. (1990).

1.3 CODING OF COLOR IN THE CORTEX

There are at least three different parallel pathways that carry the retinal signals from the lateral geniculate nucleus (LGN) to the cortex. Each projects to different regions in the primary visual cortex: The magnocellular pathway transports achromatic luminance signals, originating in parasol ganglion cells; it emerges from the magnocellular layer of LGN and terminates in sublayers 4Cα and 6 of the cortex. The parvocellular pathway, which carries L-M/M-L cone-opponent signals, projects from the parvocellular cell layer of LGN to cortical sublayers 4Cβ and 6 (Blasdel & Lund, 1983; Dacey, 2000; Henderickson, Wilson, & Ogren, 1978). The koniocellular pathway transporting the S-LM cone-opponent signal from LGN terminates in the superficial layers of the cortex, i.e. layer 1 and the cytochrome oxidase "blobs" (CO blobs) of layers 2 and 3 (Hendry & Yoshioka, 1994; Catterjee & Callaway, 2003). Besides these well-established projections, other pathways yet to be identified might transport information from LGN to the cortex (Nassi & Callaway, 2009). Receptive fields of neurons in the cortex are larger than those in the LGN (Angelucci & Sainsbury, 2006).

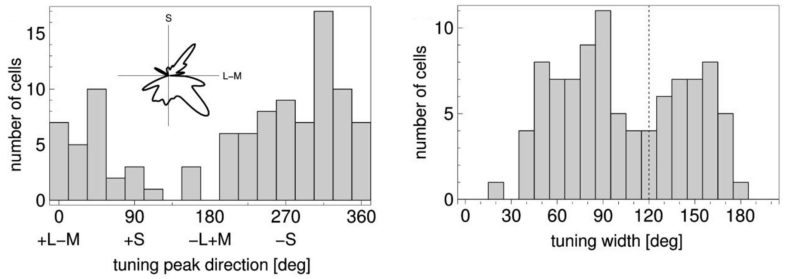


Figure 1.6: Distribution of tuning parameters of color selective neurons in primary visual cortex. (a) Distribution of tuning peaks in colorspace. (b) Distribution of tuning widths (full width at half height). Dashed vertical line in (b) indicates the location of cosine tuning. Figure reprinted with permission from Wachtler et al. (2003).

Early studies found that color selectivity of neurons in the cortex is more diverse and complex than that of LGN neurons (Thorell, De Valois, & Albrecht, 1984). Initial attempts by Livingstone and Hubel (1984) to group cells by their chromatic selectivities into distinct classes (broadband, color-opponent center-surround, color-opponent center-only and double-opponent cells) proved to be difficult and soon had to be amended (Ts'o & Gilbert, 1988; Michael, 1978b, 1978a). Systematic investigation of color selectivities by stimuli defined in a cone-opponent colorspace (cf. Section 1.2.1) by Lennie et al. (1990) revealed that chromatic preferences of neurons are largely distributed throughout the colorspace. Figure 1.5 shows the distribution of hue (azimuth) and luminance (elevation) that neurons would maximally respond to in the visual cortex (Fig. 1.5b), in contrast to the LGN (Fig. 1.5a). Subsequent studies confirmed the widely distributed nature of color preferences but reported different clusters of hue selectivities (De Valois, Cottaris, & Elfar, 2000; Hanazawa, Komatsu, & Murakami, 2000; Thorell et al., 1984; Wachtler et al., 2003; Kuriki, Sun, Ueno, Tanaka, & Cheng, 2015). The most likely reason for this is that different kinds of stimuli were used in these studies. While Lennie et al. (1990) reported clusters of hue preferences around the cardinal axes of colorspace, Wachtler et al. (2003), who used homogeneous colored full field stimuli, found almost no cells that were selective for these axes (cf. Figure 1.6b). Other studies have identified clusters corresponding to red, green, blue and yellow (De Valois et al., 2000) and also purple (Wachtler et al., 2003; Kuriki et al., 2015). The direction of the purple cluster in colorspace had a striking similarity to the opponency axis of basis functions in efficient codes for natural images (cf. Figure 1.4, T.-W. Lee et al., 2002; Wachtler et al., 2001). The shape of tuning curves of individual

neurons to color stimuli was unimodal (Lennie et al., 1990; De Valois et al., 2000; Wachtler et al., 2003). Tuning width and strength did not depend on preferred hue and the distribution of tuning widths was broad and bimodal (cf. Figure 1.6b, Wachtler et al., 2003). In summary, the distribution of chromatic preferences in primary cortex is broad, continuous but not uniform. This distributed code for color is also found at higher stages such as V2, V3 and V4 (Gegenfurtner, 2003; Gegenfurtner, Kiper, & Fenstemaker, 1996; Gegenfurtner, Kiper, & Levitt, 1997; Kiper, Fenstemaker, & Gegenfurtner, 1997; Komatsu, 1998; Komatsu, Ideura, Kaji, & Yamane, 1992). Together with the findings on distributed codes of other features and in other modalities (cf. Section 1.1.1), this suggests that a population code is a general organizational principle of the cortex. An open question is the mechanism that transforms the precortical signals into the selectivities found in primary visual cortex. Initial models assumed a linear transformation of the cone-opponent signals (Lennie et al., 1990); this model predicted cosine-shaped tuning curves. This conflicts with the range of tuning widths and with the peaks found in the bimodal distribution, which would rather suggest a non-linear combination (Wachtler et al., 2003; De Valois et al., 2000). Other response properties of neurons are also incompatible with a linear model (Horwitz, Chichilnisky, & Albright, 2005, 2007; Cottaris & De Valois, 1998). The exact nature of the non-linear transformation is still unresolved.

1.4 CONTEXTUAL EFFECTS AND PROCESSING

Context is highly important for perception, as demonstrated by various optical phenomena. Although they are often called illusions, many — if not all of them — are based on biological mechanisms that are necessary for normal vision. An example that is essential for every day vision is *color constancy*: the ability of the visual system to produce a (relatively) stable color percept under changing illuminations. This means that a cherry will consistently appear to be red both when viewed under a clear blue "cold" sky, or when a warm candlelight is the light source. It is not an easy task for the visual system to produce this stable percept, since the spectrum of the incident light from an object that will eventually reach the photoreceptors I is a combination of the reflectance of the object's surface S and the spectrum of the illuminant (E): $I(\lambda) = E(\lambda)S(\lambda)$, with λ denoting the wavelength. In order to achieve color constancy, the visual system has to discount the color of the illuminant to approximate the reflectance. When assigning colors to objects it therefore needs to take the context into account. As a result, the context can modulate the color appearance

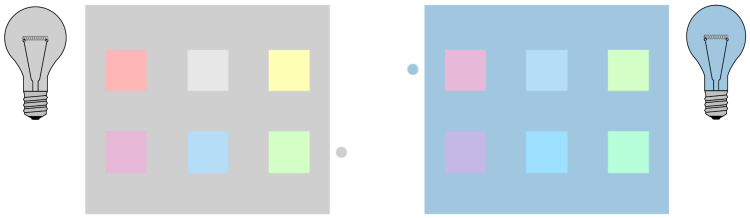


Figure 1.7: Demonstration of contextual effects in color perception. Color patches in the rows marked with a dot have the same spectral composition, i.e. the same RGB pixel values. Figure modified with permission from Wachtler et al., 2003.

of a surface so that a grayish patch looks blueish when presented against a yellow background, et vice versa.

A striking demonstration of such contextual modulation can be seen in Figure 1.7. Colored patches are presented on a neutral gray background (left) and on a chromatic background (right). The colored background introduces a hue shift that gives the impression that patches in the same row of both backgrounds have similar colors. In reality, however, only the bottom row patches of the gray background and the top row patches of the blue background (marked with a colored dot) have the same pixel RGB values, i.e. they have the same spectral composition. The blueish surround color, simulating a blueish illuminant, therefore modulates the perception of the patches in such a way that patches with different physical color values appear more similar than those whose physical color is actually the same.

Klauke and Wachtler (2015) systematically investigated such hue shifts. Human observers had to match the hue of a colored patch presented on a neutral gray surround to the hue of a patch presented on a chromatic background. Eight hues equally spaced in an isoluminant plane of the cone-opponent colorspace were presented against eight different surrounds from the same isoluminant plane. Hue shifts were measured as the difference between the matched hue and the presented stimulus hue. Each surround introduced a systematic shift that depended on the angular difference between the surround hue and the test hue. Shifts for different backgrounds were qualitatively the same but varied in strength and shape. For all backgrounds the shift was in the direction away from the hue of the surround. Wachtler et al. (2003) observed strikingly similar effects in electrophysiological recordings of color selective cells in the primary visual cortex of macaque monkeys. When full-field homogeneous

chromatic patches were presented on a chromatic background, neurons changed their tuning compared to when the patches were presented on a neutral gray background. When the collective response pattern to stimuli on chromatic backgrounds was analyzed and compared to the activation pattern of the neurons' activity on a neutral background, a hue shift very similar to that found in human could be observed. Therefore it is likely that neurons in the visual cortex contribute to color appearance and to contextual modulation effects.

1.5 OVERVIEW

The central guiding hypothesis of this thesis is that color is represented by a population code in the early visual cortex, and that such a code is suitable for further cortical processing, such as the implementation of contextual effects. As signals that reach the cortex from the retina via the LGN are encoded in a cone-opponent manner, the code undergoes a fundamental transformation. The first paper presented in Chapter 2 addresses the role of efficient coding in this transformation. By using natural images and a retina-like prefiltering, we argue that a population code is a more efficient representation of natural scenes. In Chapter 3 we simulate a population code for color together with contextual modulation to model human hue induction effects. The mechanisms underlying these effects are considered to be important for color constancy, a crucial feature of the visual system that enables stable color percepts under changing illuminations. In Chapter 4 we assess the influence of stimulus size on hue induction effects. The motivation for this study was to investigate the interplay between the origin of induction effects at the color edge and the mechanism responsible for creating a uniformly colored surface. While the former is thought to be mediated by lateral interactions, a filling in step is then necessary to create a homogeneous surface area. The results can be used to guide future hue induction models.

2

A POPULATION CODE FOR COLOR IS EFFICIENT

The one is made up of all things, and all things issue from the one.

— Heraclitus of Ephesus

2.1 SUMMARY

The coding of color information in the early visual cortex of trichromatic primates is radically different from that in the retina and the lateral geniculate nucleus (LGN). In the retina the signal is encoded via an opponent code, whereby differences between the cone signals are transmitted to the cortex via the LGN in two chromatic pathways. Studies that applied independent component analysis (ICA) on cone signals derived from natural scenes haven shown that such an opponent code is indeed an efficient way to transfer information. However, the opponency axes derived from ICA were not aligned with those used by the retinal pathways, suggesting that a more efficient representation could in theory be achieved. The encoding used by the visual system is most likely constrained by the limited number of fibers that make up the optic nerve. In the cortex, with its explosion of neurons, this constraint is gone and a reorganization of the code might lead to a more efficient representation.

Previous studies that learned efficient codes for chromatic signals in natural scenes did not take into account how cone signals are processed at the level of the retina, i.e. the opponent center surround coding and the splitting of the signal into ON and OFF pathways. Here, we mimicked the center surround filtering of the retina by filtering the cone signals derived from natural scenes with a Mexican hat like kernel. We then split the cone signals into two per channel by half wave rectification to simulate the ON and OFF chromatic pathways. Applying ICA to these signals resulted in an efficient code that was highly sparse and disperse, i.e. only a few basis functions, the analog of receptive fields of neurons, were active for a given

stimulus but always different units for different stimuli. Moreover, the chromatic preferences of the basis functions were broadly distributed in colorspace, in agreement with empirical data, which suggests that color is represented by a population code in the visual cortex, and that this representation is efficient.

2.2 CONTRIBUTIONS

The contributions of the authors Christian Johannes Kellner (CJK) and Thomas Wachtler (TW) are as follows: CJK and TW designed the study. CJK implemented and performed the simulation. CJK and TW wrote the manuscript. First results were presented at the *Bernstein Conference 2011* in Freiburg. The manuscript was published in *Frontiers in Psychology* as:

Kellner, C. J. & Wachtler, T. (2013). A distributed code for color in natural scenes derived from center-surround filtered cone signals. *Front. Psychol. 4*(661). doi: 10.3389/fpsyg.2013.00661.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 27 September 2013
doi: 10.3389/fpsyg.2013.00661

A distributed code for color in natural scenes derived from center-surround filtered cone signals

***Christian J. Kellner*[1,2]* and *Thomas Wachtler*[1,2,3]**

[1] *Department of Biology II, Ludwig-Maximilians-Universität München, Martinsried, Germany*
[2] *Graduate School of Systemic Neurosciences, Ludwig-Maximilians-Universität München, Martinsried, Germany*
[3] *Bernstein Center for Computational Neuroscience, Ludwig-Maximilians-Universität München, Martinsried, Germany*

Edited by:
Galina Paramei, Liverpool Hope University, UK

Reviewed by:
Sérgio M. C. Nascimento, University of Minho, Portugal
C. Alejandro Párraga, Universitat Autònoma de Barcelona, Spain

***Correspondence:**
Christian J. Kellner, Department of Biology II, Computational Neuroscience, Ludwig-Maximilians-Universität München, 82152 Martinsried, Germany
e-mail: kellner@bio.lmu.de

In the retina of trichromatic primates, chromatic information is encoded in an opponent fashion and transmitted to the lateral geniculate nucleus (LGN) and visual cortex via parallel pathways. Chromatic selectivities of neurons in the LGN form two separate clusters, corresponding to two classes of cone opponency. In the visual cortex, however, the chromatic selectivities are more distributed, which is in accordance with a population code for color. Previous studies of cone signals in natural scenes typically found opponent codes with chromatic selectivities corresponding to two directions in color space. Here we investigated how the non-linear spatio-chromatic filtering in the retina influences the encoding of color signals. Cone signals were derived from hyper-spectral images of natural scenes and preprocessed by center-surround filtering and rectification, resulting in parallel ON and OFF channels. Independent Component Analysis (ICA) on these signals yielded a highly sparse code with basis functions that showed spatio-chromatic selectivities. In contrast to previous analyses of linear transformations of cone signals, chromatic selectivities were not restricted to two main chromatic axes, but were more continuously distributed in color space, similar to the population code of color in the early visual cortex. Our results indicate that spatio-chromatic processing in the retina leads to a more distributed and more efficient code for natural scenes.

Keywords: color vision, visual cortex, sparse coding, natural image statistics, population code, efficient coding

1. INTRODUCTION

In the retina of trichromatic primates, spatio-chromatic processing of signals from long (L), medium (M), and short (S) wavelength selective cones by ON and OFF bipolar and ganglion cells with center-surround receptive fields leads to parallel pathways that carry both spatial and chromatic information to the lateral geniculate nucleus (LGN) and visual cortex (Lee, 2011). Chromatic signals are carried in pathways encoding differences between S cone signals and the signals of L and M cones, or differences between L and M cone signals, respectively (Mollon, 1989; Dacey, 2000; Solomon and Lennie, 2007; Lee et al., 2010). While the first, phylogenetically older pathway has low spatial selectivity and is thought to be specifically concerned with color information, the second pathway carries both spatial and chromatic information (Boycott and Wässle, 1999; Martin et al., 2011). Midget retinal ganglion cells have spatially center-surround receptive fields and in the fovea achieve their color opponency by antagonistic processing of signals from a single cone in the center and several cones in the surround of their receptive field. There is evidence for functional cone-type specificity beyond that arising from a single-cone center, but different studies have arrived at different conclusions (Reid and Shapley, 1992, 2002; Lee et al., 1998; Martin et al., 2001; Buzás et al., 2006; Field et al., 2010; Crook et al., 2011; Martin et al., 2011; Lee et al., 2012), and the question of the degree to which cone-type specific wiring contributes to midget receptive fields remains open.

Corresponding to the parallel pathways in the retina, color selectivities in the LGN cluster around the two cardinal axes of cone opponency (Derrington et al., 1984). In visual cortex, however, the representation of color is different. Chromatic information is encoded cortically by both opponent and non-opponent neurons (Lennie et al., 1990; Wachtler et al., 2003). Moreover, the preferences of cortical color-selective neurons are not restricted to two main axes of opponency, but are more distributed (Lennie et al., 1990), indicating a population code for color (Wachtler et al., 2003). The transformation from coding along cone opponency axes to a distributed representation in the cortex is not well understood, but according to the theory of efficient coding (Barlow, 1961, 2001) one hypothesis would be that this code is in some sense more efficient.

Previous studies investigating efficient codes for color in natural scenes have used independent component analysis (ICA), a method for finding a linear transformation that makes the resulting outputs as statistically independent as possible (Jutten and Herault, 1991; Bell and Sejnowski, 1997). Analyses of chromatic natural images using ICA revealed that opponent codes are efficient to encode natural color stimuli. Typically, in these studies two main types of chromatic selectivity were found (Hoyer and Hyvärinen, 2000; Wachtler et al., 2001; Doi et al., 2003), which qualitatively resembled more the representation in retina and LGN than the coding properties in the visual cortex. While the discrepancies can be explained in part by the stimuli used in the

experiments to determine color tuning in visual cortex (Caywood et al., 2004), broadly distributed color selectivities have also been found with other types of stimuli (Wachtler et al., 2003).

A reason for the lack of insights about the nature of the distributed cortical representation of color from previous studies could be that the assumed model of a linear transformation of cone signals was not appropriate. Comparison of efficient codes found by methods like ICA with properties of neurons in the visual system requires that visual processing can be approximated as a linear transformation of the cone signals. However, the spatio-chromatic processing in the retina transforms the cone signals in fundamentally nonlinear ways and a linear model may not be adequate. Moreover, spatial center-surround filtering as observed in retinal neurons removes much of the spatial correlations between signals of neighboring cones (Wachtler et al., 2007), which may enhance the relative contribution of chromatic variation. It is further conceivable, given the limited number of fibers in the optic nerve as compared to both the number of receptors in the retina and the number of neurons in primary visual cortex, that retinal processing is subject to different constraints and coding objectives than the representation in the visual cortex (Lee et al., 2002). Under this assumption, retinal processing could be considered as a preprocessing step separate from cortical processing (Doi et al., 2003), and it would be appropriate to perform the analyses not on the cone signals, but on the output signals of the retina. Here we investigated the consequences of nonlinear spatio-chromatic filtering in the retina for the efficient coding of chromatic information in natural scenes. We modeled the center-surround processing in the retina to obtain estimates of the signals in the different parallel retinal pathways carrying chromatic information and analyzed these signals by performing ICA.

2. MATERIALS AND METHODS

2.1. IMAGE BASIS

As image basis for the main analysis we used eight images from the hyper-spectral image database by Párraga et al. (1998) that were previously used to study efficient codes for color (Wachtler et al., 2001; Lee et al., 2002). We used the same set of images in order to make our results directly comparable with these studies. The images were recorded outdoors around Bristol, UK, under stable weather and lighting conditions (Párraga et al., 1998). In additional analyses, we used other images from the Párraga et al. dataset as well as from other datasets (see below). Images were 256 × 256 pixels in size, with each pixel subtending 0.056 × 0.056 degree of visual angle. Pixels corresponded to radiance values in 31 wavebands between 400 nm and 700 nm. Radiance values were derived from the raw data with the code provided by Párraga et al. (1998). In all scenes a Kodak GrayCard reflectance standard was present; the corresponding picture areas were ignored during analysis.

To control for potential misalignment between the color planes in the hyper-spectral images due to the relatively long acquisition time (Párraga et al., 1998), we estimated the drift between image planes by 2-d cross-correlation. In most cases, the misalignment was zero, and non-zero misalignments appeared unsystematic, with a maximum shift of 2 pixels. Repeating the analysis with images where these shifts had been corrected did not alter the findings. As an additional control we used the four images of the hyper-spectral dataset of Foster and Nascimento (Nascimento et al., 2002; Foster et al., 2006) that showed natural scenes. The individual images were recorded within 15 s, which largely excluded any misalignment of wavelength planes. The results of this analysis, as well as those of an analysis with all the images in this dataset that were larger than 600 × 600 pixels, did not change any of the findings.

2.2. IMAGE FILTERING

To take into account the spatio-chromatic filtering by the retina, three main processing steps were modeled: (a) transduction of photons to neural signals by the photoreceptors, (b) center-surround integration of cone signals, and (c) splitting of the signals in ON and OFF pathways, mediated by bipolar cells with center-surround receptive fields. **Figure 1** provides an overview of the entire filtering process.

To obtain cone excitations from natural scenes, for each of the images we computed the dot products of the pixel spectra with the vectors of cone sensitivities, resulting in a 256 × 256 × 3 matrix. Human cone sensitivity estimates were taken from Stockman et al. (1993). The center-surround integration stage was modeled by

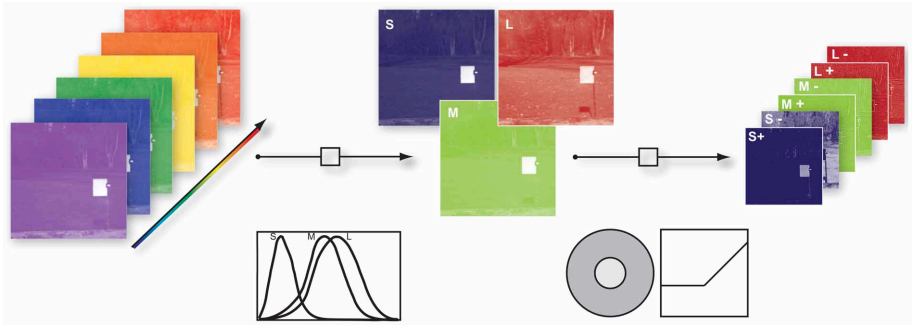


FIGURE 1 | Illustration of the image filtering process. (1) Illustration of hyper-spectral image. Each image consisted of 31 image planes corresponding to radiance within wavebands between 400 nm and 700 nm (2) S-, M-, and L- photoreceptor cone activation were derived via their specific sensitivities. (3) S, M, and L center ON and OFF channels were obtained by convolution of the cone activations with a Mexican hat-like spatial filter, simulating the center-surround receptive fields of the retina, and splitting the result into positive and negative parts.

convolution of the image with a Mexican hat-like kernel. We used an approximation of a Difference-of-Gaussian, consisting of a 3×3 pixel matrix with a value of 1 at the center, -0.15 at the center pixel of each edge, and -0.1 at the corner pixels. This filtering assumed that total weights of center and surround are balanced, and that each pixel of each pixel plane represents a ganglion cell with a single cone in the center. We assumed that the surround consists of all three cones exerting an equal contribution at each location (mixed L,M,S surround), but tested other configurations as well. ON and OFF signal channels were generated by half-wave rectification on the filter outputs and their sign-inverted counterparts, respectively. After the rectification procedure we log-transformed every channel to mimick a compressive response function. Since the rectification step introduced zero values into the data and the natural logarithm diverges at zero we added a dynamic offset to the channel. The offset was chosen such that all channels had the same dynamic range.

For the analysis, 7×7 patches were selected randomly from the prefiltered data. ON and OFF pixel planes for all 3 cone classes were interleaved at each pixel. The resulting dimensionality of a single input data sample was thus $7 \times 7 \times 3 \times 2 = 294$.

2.3. ICA

ICA was proposed as a solution to the blind source separation problem and has been applied in various studies (Bell and Sejnowski, 1997; Wachtler et al., 2001; Lewicki, 2002) to learn efficient codes for visual stimuli. The ICA model assumes a linear mixture of statistically independent sources $\boldsymbol{s}$ (also often called causes), which is observed via a number of sensors. If no additive sensor noise is assumed, the problem can be written as:

$$\boldsymbol{x} = \boldsymbol{A}\,\boldsymbol{s} \tag{1}$$

Note that neither the sources $\boldsymbol{s}$ nor the mixing matrix $\boldsymbol{A}$ are known. The goal of ICA is to recover the sources by adapting $\boldsymbol{A}$ such that the resulting signals are as statistically independent as possible.

Once $\boldsymbol{A}$ has been inferred, the source can simply be uncovered by solving for $\boldsymbol{s}$:

$$\boldsymbol{s} = \boldsymbol{A}^{-1}\,\boldsymbol{x} = \boldsymbol{W}\,\boldsymbol{x} \tag{2}$$

The columns of $\boldsymbol{A}$ are usually called the basis functions and the rows of $\boldsymbol{W}$ are called the filters.

Here we used the approach by Lee and Lewicki (2000) with the learning rule for $\boldsymbol{A}$ given by:

$$\Delta\boldsymbol{A} \propto \boldsymbol{A}\boldsymbol{A}^T \frac{\partial}{\partial \boldsymbol{A}} \log\, p(\boldsymbol{x} \mid \boldsymbol{A}) = -\boldsymbol{A}(z(\boldsymbol{s})\boldsymbol{s}^T - \boldsymbol{I}). \tag{3}$$

The individual terms are the identity matrix $\boldsymbol{I}$, the transpose of the sources $\boldsymbol{s}^T$ and $z(\boldsymbol{s}) = \frac{\partial\, \log\, p(\boldsymbol{s})}{\partial\, \boldsymbol{s}}$. The prior source distributions were modeled using the exponential power distribution (also known as the generalized Gaussian or generalized Laplacian). The simple form is:

$$p(s_i) \propto e^{-\frac{1}{2}|s_i|^{q_i}} \tag{4}$$

The kurtosis can be controlled by varying q_i and thus platykurtic, leptokurtic, and Gaussian distributions can be modeled. We used $q_i = \frac{2}{1+\beta_i}$ and estimated β_i during learning. Therefore no additional assumption about the exact distribution of the sources were made a priori. As β_i becomes bigger, the distribution becomes more leptokurtic and the resulting code more sparse, meaning that the source coefficients are mostly close to zero.

The mixing matrix $\boldsymbol{A}$ was estimated in 100.000 iterations. It was initialized with Gaussian distributed random values and all priors were set to Gaussian densities. After every 400 iterations new input data were sampled by drawing 5000 patches randomly from each of the eight pictures and β_i was re-estimated. All samples were centered and rescaled to have zero mean and unit variance. The stepsize was adjusted at iteration points 1000, 5000, 10000, 30000, 70000, to 0.02, 0.01, 0.005, 0.002, 0.001 and 0.0001 respectively. In order to accelerate the learning process, the algorithm was ported to the CUDA parallel computing architecture and run on a NVIDIA Tesla M2090 graphics processor.

2.4. ANALYSIS OF THE RESULTS

2.4.1. Reverse correlation - activation triggered averages

After learning, the mixing matrix $\boldsymbol{A}$ and the unmixing matrix $\boldsymbol{W}$ were adapted to the preprocessed data. Due to the non-invertible nonlinear filtering, the result was not a simple linear unmixing of LMS signals. Therefore, we used a reverse-correlation approach to illustrate the resulting filters: Source activations for each post-filtered patch were used as weights for the corresponding original patch in LMS-color space. By averaging over all weighted original patches we computed the activation triggered average (ATA), i.e., the average patch that would elicit the maximal response for a single basis function. The details of this procedure are as follows:

When the individual filters w_r (rows of $\boldsymbol{W}$) are used to perform the unmixing of the data, each individual source coefficient s_r is a direct measure of the response of the filter w_r to a given data sample x_r. In our case the data samples were the preprocessed patches at $p(\boldsymbol{i})$, where the vector $\boldsymbol{i} = (x, y, e)$ specifies the patch position (x, y) in the preprocessed image e. Using the transformation between the preprocessed patches $p_k(\boldsymbol{i})$ and the patches in LMS space $\hat{p}(\boldsymbol{i})$, we can then calculate the average original patch ATA_r that the individual filters w_r best respond to by using the source coefficient derived from $p(\boldsymbol{i})$ as a weight for $\hat{p}(\boldsymbol{i})$ and then averaging over all available $\hat{p}(\boldsymbol{i})$.

We therefore generated all possible 7×7 patches from each of the eight preprocessed images used for analysis ($N = 8 * 61504 = 492032$). The source activations were then estimated using equation (2). To eliminate noisy contribution of source coefficients with a very low absolute activation, i.e., source activations around zero, we fitted the source activation with an exponential power distribution. When the mean of the fit was close to zero and the distribution was leptokurtic we used the cumulative distribution function $F(x)$ to discard 95% of all the source activation around the peak (see **Figure 2**). Thus for each basis function and each patch in every image we computed a weight $\alpha_r(\boldsymbol{i})$:

$$\alpha_r(\boldsymbol{i}) = \begin{cases} 0 & \text{if } F(x) > 0.25 \wedge F(x) < 0.975 \\ s_r(\boldsymbol{i}) & \text{otherwise} \end{cases} \tag{5}$$

To calculate the patch in LMS-space that each basis function would maximally respond to (the ATA), we weighted each original

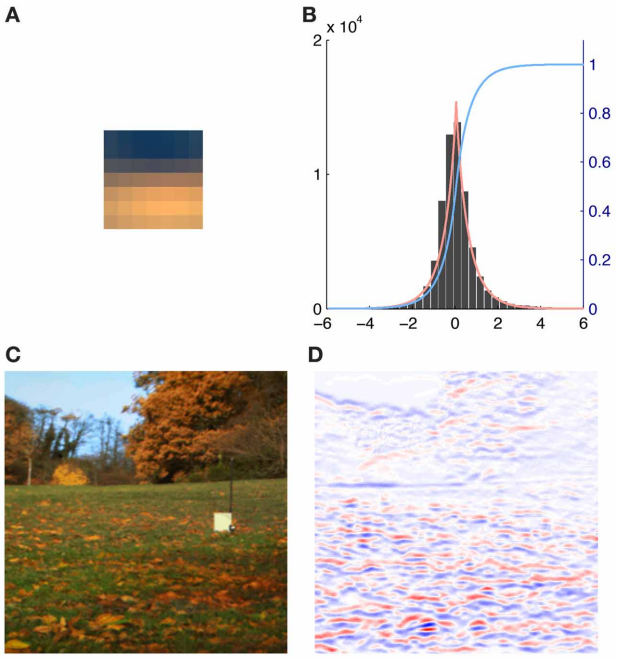


FIGURE 2 | Illustration of the contribution of a single basis function (number 116) to an image together with the exponential power distribution fit of the corresponding source activations. (A) Activation triggered average of basis function 116. **(B)** Histogram of the source coefficient values showing a typical leptokurtic distribution with a peak around zero and heavy tails. The fit of the exponential power distribution to the data (β: 1.312, μ: 0.001, σ: 0.753) is shown in red and the cumulative distribution function $F(x)$ is plotted in blue. **(C)** RGB rendering of the original image. **(D)** Contributions to the image, red indicates high positive, blue high negative and white no contributions of basis function 116.

patch $\hat{p}(\boldsymbol{i})$ with the corresponding weight $\alpha_r(\boldsymbol{i})$ that was calculated earlier and averaged over the result:

$$ATA_r = \frac{1}{N}\sum_{k=1}^{N} \alpha_r(k)\,\hat{p}(k) \tag{6}$$

We verified the plausibility of this approach by comparing the ATA_r with the basis functions for the analysis of LMS input without preprocessing as in Lee et al. (2002). ATAs tended to be slightly less saturated than the corresponding basis functions but otherwise they resembled each other in color preference and spatial structure almost completely.

2.4.2. Plotting of the results

We displayed the ATA_r as shown in **Figure 3** with the method used by Ruderman et al. (1998). The L, M, and S components of each pixel of the patches were first normalized to values between 0 and 1 and then plotted as red (R), green (G) and blue (B) values. This gives a pseudo-color representation of relative cone excitations that is qualitatively similar to a true-color rendering. Therefore spatial as well as chromatic structure can be observed. To further illustrate the chromatic properties, each pixel was plotted as a point in a cone-opponent color space (Wachtler et al., 2001), where x values were computed as $x = L - M$, y values $y = S - (L + M)/2$ and the z values as $z = L + M$. When x, y, z values are converted to a spherical representation r, θ, ϕ, the azimuth angle ϕ is a direct measure of the hue of a given point while the radial distance r indicates its chroma and the elevation θ specifies the luminance. For plotting points we used a projection of the z-axis onto an isoluminant plane. Luminance information can still be inferred from the brightness of the individual points.

2.4.3. Directions in color space

Once transferred to the cone-opponent color space, the chromatic characteristics of each patch could be quantitatively studied. To quantify the degree of opponency of individual patches, i.e., whether the pixel selectivities were roughly aligned in color space, we performed principal component analysis on the color space coordinates of all the pixels in each patch. The highest eigenvalue was used as an estimate of the strength of opponency, and the eigenvectors were used to estimate the directions of opponency. Additionally, the average color preference for a given patch was calculated by the center of mass of all points.

To quantify how uniformly a set of directions O were spread out in color space we calculated the Kullback-Leibler-Divergence (D_{KL}) from a uniform distribution U with the same number of directions as O: $D_U(O) = D_{KL}(U \parallel O) = \sum_i ln(\frac{U(i)}{O(i)})U(i)$. The higher the value of this measure, the higher the divergence from uniformity. This measure can then be used to compare different sets of directions O_n derived from different ATAs.

2.5. SPARSENESS CHARACTERIZATION

To quantify the sparseness of the resulting code and thus its efficiency, we used the criteria proposed by Willmore and Tolhurst (2001): The mean lifetime kurtosis $\overline{K_L}$, the population kurtosis $\overline{K_P}$ and the dispersal of the learned code. Both kurtosis values were computed via the standard kurtosis measure. The lifetime kurtosis K_L of the response, i.e the source activation of a single component is a measure of how active this component is across all stimuli. The population kurtosis K_P quantifies how many filters are active to encode a single stimulus. A high average K_P-value means that only a small number of available filters are active for any given input. The dispersal of the code is a measure of the contribution of each filter to the encoding of the data. It is based on measuring the variance of the response of a filter to the image data. For a given code the standard deviation of all filters is estimated for each image and then normalized to the highest value and sorted according to their normalized standard deviation. In a compact code only a few filters encode the majority of the total variance of the data so the relative standard deviation of only a few filters will be high (close to one) and close to zero for all others. In a more dispersed code where individual filters have all higher contributions to the data, the relative standard deviation will be higher for all filters.

3. RESULTS

3.1. SPATIO-CHROMATIC STRUCTURE

ATAs for all 294 basis functions are shown in **Figure 3**. ATAs are sorted according to the L^2 norm of the corresponding basis functions (see also **Figure 4**). The L^2 norm can be used as a

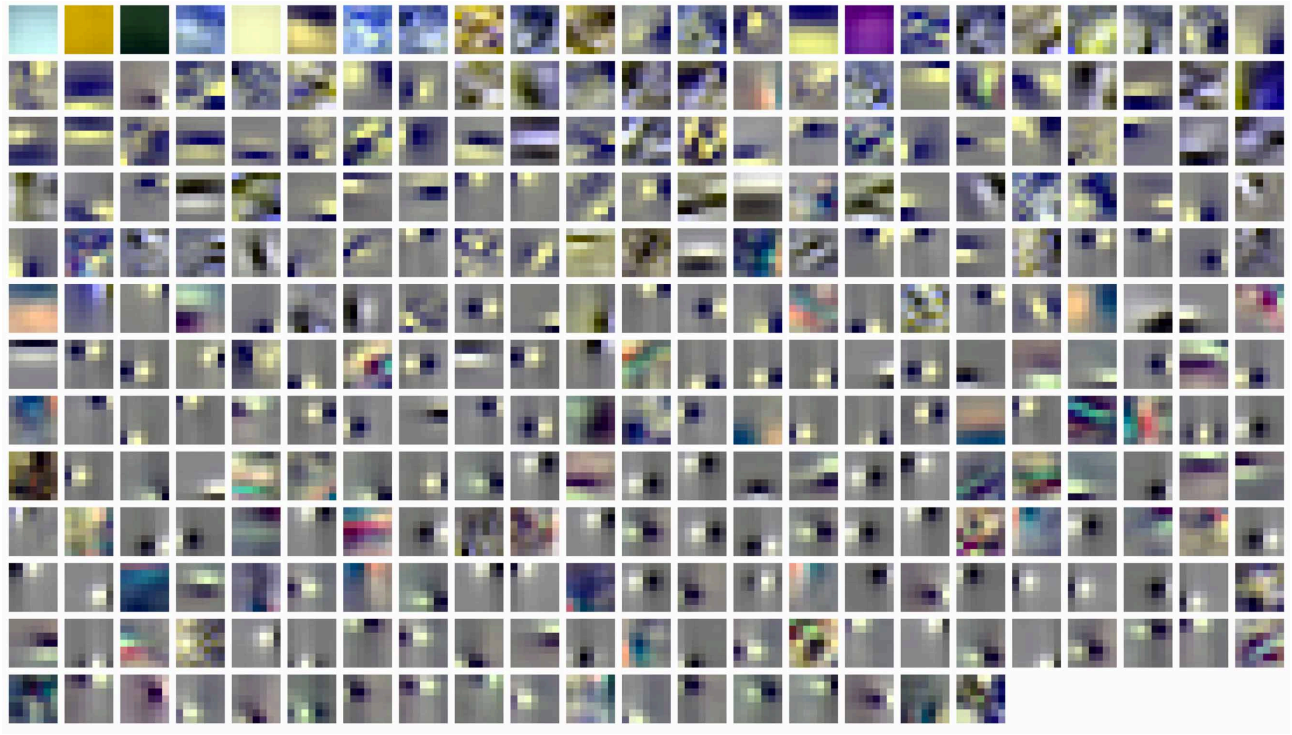


FIGURE 3 | Spatio-chromatic structure of the activation triggered averages for all 294 learned ICA basis functions. R, G, and B pixel values for each 7x7 patch correspond to L, M, and S cone excitations that were derived from individual basis functions using a reverse correlation based approach.

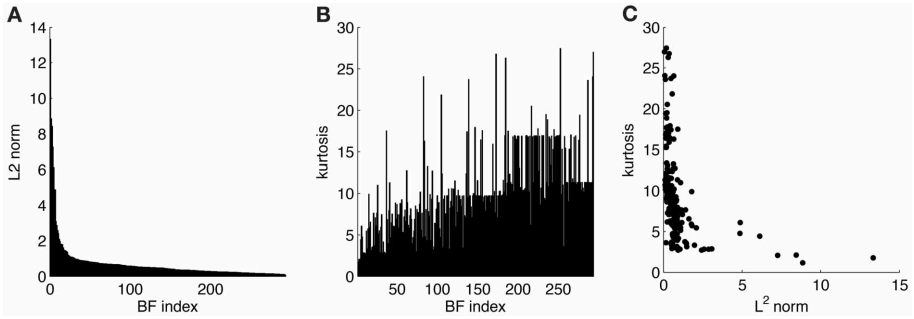


FIGURE 4 | Statistics for learned ICA basis functions. (A) Histogram of L^2 norms for every individual basis function after sorting. **(B)** Lifetime kurtosis of the source coefficients for each basis function (sorted as in **A**). **(C)** Lifetime kurtosis vs. the L^2 norm.

measure of the relative contribution of each basis function to the data. The chromaticities of the individual pixels of each ATA plotted in the cone-opponent color space are shown in **Figure 5**. Visual inspection of **Figures 3**, **5** suggests that almost all ATAs can be divided in three major categories: homogeneous chromatic, color-opponent, and achromatic. Homogeneous chromatic ATAs have a large L^2 norm, no defined spatial structure, and are highly selective for one color. Most non-homogeneous but chromatic ATAs were color-opponent, i.e., the pixel chromaticities, when plotted in color space, were all clustered along a line and most often also crossed into opposing quadrants. Their spatial structure was both localized and oriented, i.e., they encoded chromatic edges (cf. Wachtler et al., 2001). A small number of non-homogenous chromatic ATAs were less strongly opponent with their pixel values more scattered in color space. There was no substantial correlation between the L^2 norm of the basis function and the degree of opponency ($r = 0.1$). The achromatic ATAs, encoding luminance edges, had mid- to low-range contributions. This is a notable difference to previous findings (Wachtler et al., 2001; Lee et al., 2002), where many achromatic basis functions with high L^2 norm were found (see below).

3.2. DISTRIBUTION OF COLOR PREFERENCES

To illustrate the overall color preferences of ATAs, we computed the center of mass of all pixels for a single ATA. The resulting positions are plotted in **Figure 6A**. Additionally, the direction of

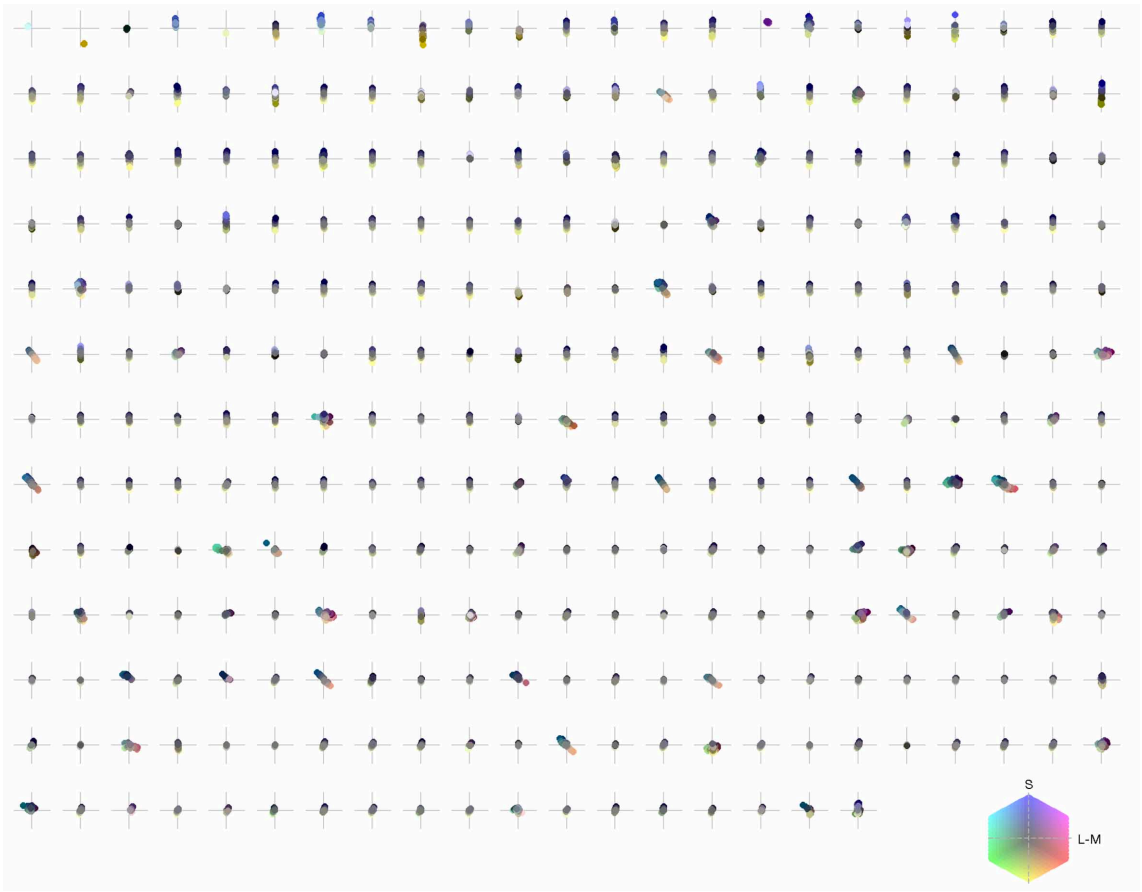


FIGURE 5 | Chromaticities of activation triggered averages: each pixel of every individual patch is projected onto an isoluminant plane in a cone-opponent color space. The horizontal axis is defined as L − M and the vertical axis as S − (L + M)/2. Luminance values correspond to brightness of pixels.

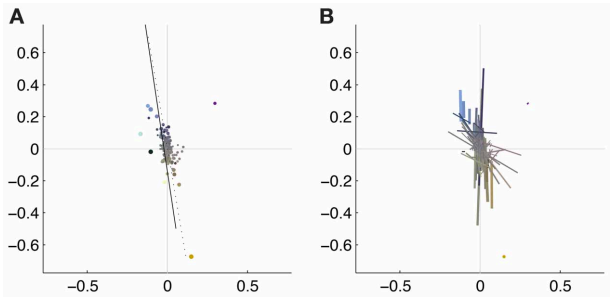


FIGURE 6 | Coverage of color space by basis functions. The thickness of the each dot **(A)** or line **(B)** is scaled with the L^2 norm of the basis function, the length of the line corresponds to the eigenvalue obtained from the PCA of the pixels chromaticities (cf. **Figure 5**). The position of the dots and the midpoint of the line indicate the center of gravity. Solid line: daylight axis; Dotted line: direction of first principal component of center of mass distribution.

largest variation around the center of mass position is shown in **Figure 6B**. The center of mass positions were all densely clustered around the origin, indicating relatively weakly pronounced selectivities, with the exception of the homogeneous ATAs, which were more eccentric. All points together formed a distribution that was strongly elongated along a certain direction in color space. This direction, as estimated by the first principal component of the distribution, had an angle of 101.6 degrees in this color space. This matches closely the perceptual "yellow"-"blue" line, approximated by the line between the loci of monochromatic blue light of 476 nm and monochromatic yellow light of 576 nm (Mollon, 2006), which also lies close to the line of natural daylight variation and has an angle of 98.5 degrees in this color space.

Figure 7 shows the distribution of color preferences across directions in color space. In contrast to previous results obtained without spatio-chromatic preprocessing (Wachtler et al., 2001; Lee et al., 2002), color preferences were spread around the entire color space. However, the distribution was not uniform but showed several regions of higher density. One of these regions was around 90 degrees, with pixel chromaticities varying between light-blue and dark-yellow. This corresponds to a modulation of values along a plane defined by S-cone and luminance variation. Many ATAs with this chromatic signature had localized and oriented spatial features that qualitatively resembled the structure of the basis functions found for natural gray-scale images (Olshausen and Field, 1996; Bell and Sejnowski, 1997) and the achromatic basis functions for L-,M-, and S-cone activations

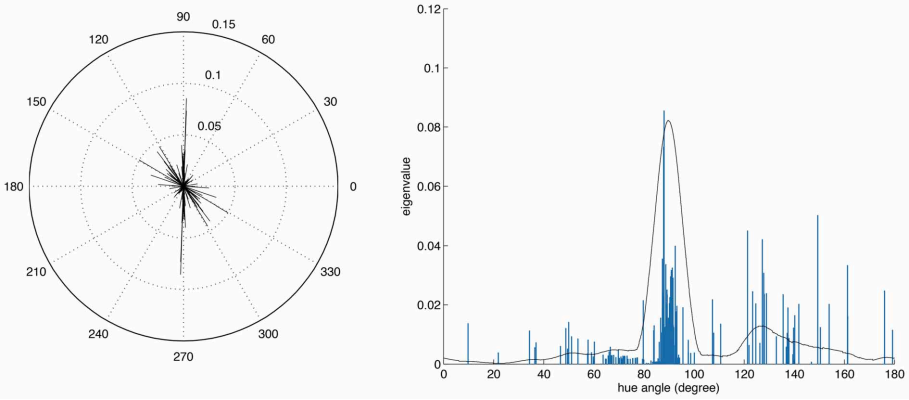


FIGURE 7 | Distributions of opponency directions in color space. Each line in the polar and linear plot indicates the direction of the opponency axis for a single ATA or basis function. The lengths of the lines indicate the spread of pixels along the opponency axis. Since most of the color-opponency axes passed through a region very close to the origin into the opposing quadrant, we only show the range between 0 and 180 degrees in the linear plot.

(Wachtler et al., 2001; Lee et al., 2002). The second region with higher density was around 130 degrees, which corresponds to an opponency axis between orange and teal. These regions correspond to the two opponency axes found in previous studies (Wachtler et al., 2001; Lee et al., 2002). Another more densely covered region appeared in the first quadrant around 65–80 degrees, and the region around 10–30 degrees, was the least densely covered area. Apart from these modulations in density, the directions of color-opponent axes were widely distributed in color space with a divergence from uniformity $D_U(O)$ of 16.65, compared to the code obtained from pure LMS cone activation (Wachtler et al., 2001; Lee et al., 2002) which had a $D_U(O)$ value of 25.58.

To further analyze the chromatic properties we determined for each ATA the color tuning of the pixel with the maximal absolute value. This is comparable to estimating the color preference for small colored spots. By using this measure the directions of color preference were even more uniformly distributed in color space, with a $D_U(O)$ of 2.59 for the filtered data as compared to 21.7 for LMS data.

3.3. CODING EFFICIENCY

Coding efficiency was originally understood in terms of redundancy reduction. Under this assumption a code is efficient if it reduces the mutual information between components, i.e., the information encoded among a group of neurons would be reduced as much as possible. Another measure of coding efficiency especially when dealing with a large number of encoding neurons, such as in the cortex, is the sparseness of the code, i.e., how many neurons of all that are available are used to encode a specific stimulus.

A quantitative study of redundancy reduction efficacies of different linear filtering algorithms was done extensively by Eichhorn et al. (2009). Multi-information reduction, Average Log-Loss and rate-distortion curves were used as evaluation criteria for various algorithms like ICA and Principal Component Analysis (PCA), which were all compared to a random decorrelation method that served as baseline. We used the source code provided with the paper and adapted it to process our prefiltered and rectified cone signal data. Even though we kept the changes to a minimum in order to stay as close to the original analysis, it was not possible to use the NPL entropy estimator for the filtered data due to numerical instabilities. The reason for this most likely is that the distribution of the data after our preprocessing does not fit with the model assumptions of the NPL entropy estimator. Therefore, we used the Gaussian upper entropy bound (Bethge, 2006). Our results are thus not directly comparable to those in Eichhorn et al. (2009). Nevertheless the absolute multi-information reduction with respect to the random decorrelation transform (RND) was one order of magnitude better for ICA than for PCA, namely -0.4640 ± 0.0058 bits/component (ICA) and -0.0460 ± 0.0013 bits/component (PCA). The relative reduction in multi-information (cf. Table 1. in Eichhorn et al., 2009) compared to RND was 0.42 ± 0.01 percent for ICA, and 0.04 ± 0.00 percent for PCA. The Average Log-Loss (ALL), as a measure of how well the density model of the specific transformation matches the actual data and differences correspond to coding cost (Eichhorn et al., 2009). The difference for PCA-RND was -0.0481 ± 0.001 bits/component, for SSD-RND (spherically symmetric density) it was -0.2429 ± 0.0001 bits/component, and for ICA it was -0.4206 ± 0.0036 bits/component. Thus, in our case ICA performed best, i.e., it had the smallest ALL value.

To estimate the sparseness of the learned representation we computed the lifetime kurtosis K_L of individual units and the population kurtosis K_P (cf. Willmore and Tolhurst, 2001). **Figure 8** shows histograms for source coefficients together with their estimated lifetime kurtosis. All source densities are highly sparse (i.e., leptokurtic), with a pronounced acute peak at zero and heavy tails. The mean lifetime kurtosis $\overline{K_L}$ over all units was 10.29, which means individual units were silent for almost all inputs but very strongly activated for specific input features. The mean population kurtosis over all inputs was 12.67, i.e., only a small subset of available neurons were active for any given input. In addition to lifeline and population kurtosis, the dispersal is an important measurement for the sparseness of the code. This measure, based on the standard deviation of the responses, quantifies the relative coding contribution of each filter (derived from the

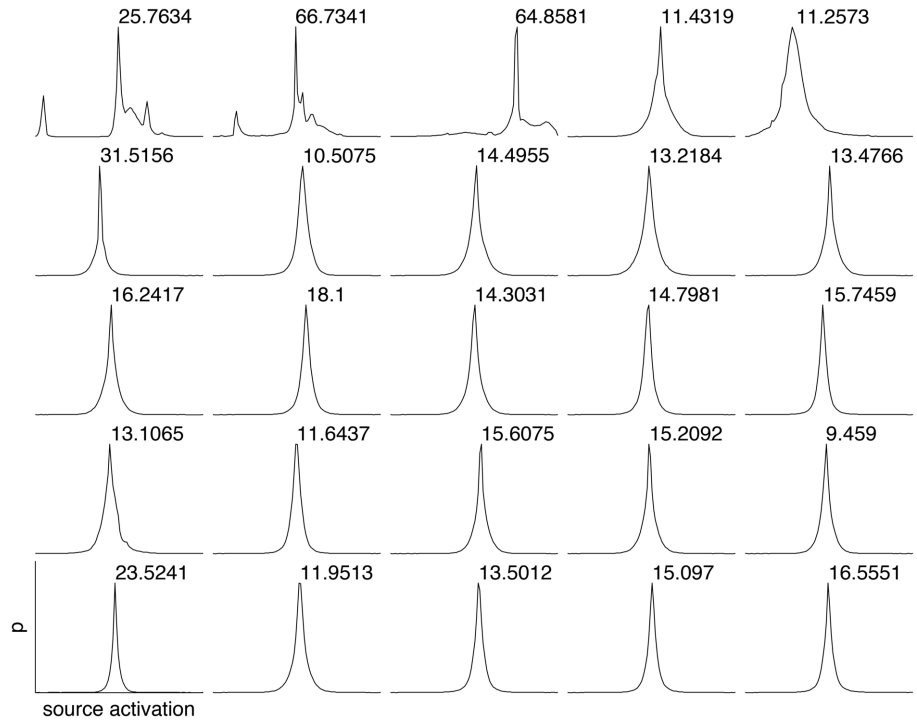


FIGURE 8 | Histograms of the source coefficient values for the first 25 basis functions and the standard kurtosis derived from the source coefficients.

basis function) to the image data. **Figure 9** shows the dispersal for ICA, and as a comparison for PCA, which is an example of a compact code. PCA was done on the same preprocessed cone activation data. For ICA the decrease in coding contribution is close to linear, while for PCA it is exponential. This confirms that PCA is a much more compact code, i.e., only a few filters are used to encode the majority of the data, while in the ICA case most of the filters have a high contribution. Overall, these results show that, in accordance with previous studies (Lee et al., 2002), the obtained representation for the preprocessed images was a highly disperse and sparse, i.e., statistically efficient, code, although sparseness itself was not an enforced constraint during learning.

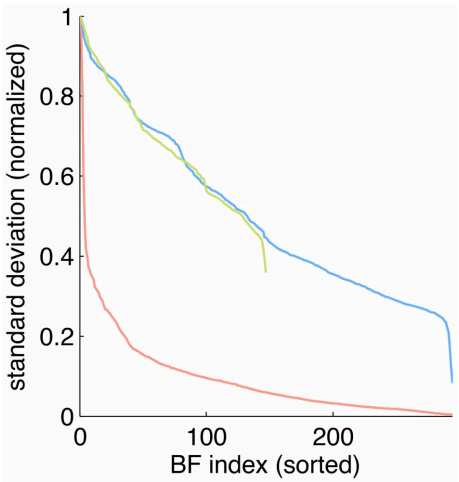


FIGURE 9 | Dispersal of different codes: ICA on preprocessed data (blue), PCA on preprocessed data (red) and ICA on LMS images (green). For each code, the mean of the per image response standard deviation normalized to the largest value is plotted for each basis function. Note that for unfiltered LMS images the dimensionality of the data is lower, yielding only 147 basis functions.

4. DISCUSSION

We investigated the consequences of nonlinear spatio-chromatic filtering similar to the processing in the retina, including the splitting into parallel ON and OFF color-opponent channels, for the learning of efficient codes from responses to natural scenes. Compared to the results of previous studies where ICA was used to learn efficient codes directly from LMS cone activations of natural images (Wachtler et al., 2001; Lee et al., 2002), chromatic preferences obtained from opponent signals were more broadly distributed in color space. A continuous distribution is in better accordance with experimental data (Lennie et al., 1990; Wachtler et al., 2003) than the strong clustering into three chromatic types observed previously. Additionally, it is also in closer correspondence with precortical encoding of color. The filtering we applied mimicks the effect of center-surround receptive fields of retinal bipolar and ganglion cells, which removes redundancy both in the spatial and the spectral domain. In previous studies, whitening had been applied in a linear preprocessing stage before ICA. However, to estimate the results, this pre-filtering had to be taken into account by adding a corresponding linear transformation. In our analysis, such a direct compensation would not have

been possible because the preprocessing stage was a nonlinear transformation. To represent the resulting components, we therefore used a reverse correlation technique to obtain a single-stage linear transformation representing the effective linear component of the multi-stage nonlinear filtering. A further effect of the preprocessing was the representation of the signals in the higher-dimensional space of six rectified opponent channels. This representation may have facilitated the distinction of features (Schölkopf and Smola, 2002). Similarly, the parallel channels in the retina and LGN provide such a high-dimensional representation, which might be exploited by cortical learning mechanisms.

The set of natural images was chosen to be the same as in Wachtler et al. (2001); Lee et al. (2002) to enable direct comparison of the results. These images were initially chosen to include a variety of scenes recorded outdoors and under different illuminations. To exclude that our results were an artifact of the specific choice of images, we repeated the analysis including all outdoor images contained in the Párraga et al. (1998) dataset. The result was again a broad distribution of chromatic preferences with a divergence from uniformity of 15.91, compared to a value of 28.64 obtained for these images without pre-filtering. In addition, we ran an analysis using a larger patch size of 10×10 pixel. This resulted in a spread of selectivities that was even more broad, with a divergence from uniformity of 11.09 compared to 24.27 obtained without pre-filtering.

For the center-surround spatio-chromatic filtering we used filters with a surround composed of equal contributions of all cone types. Often, the center-surround processing in the retina is likened to a whitening stage that removes second-order dependencies (Doi et al., 2003). Whitening filters for LMS images typically have also a cone-type specific center-surround structure. We repeated the analysis using whitening filters for the preprocessing. The color preferences of the resulting ICA ATAs were strongly clustered around a single region in color space, which is not in line with the observed color preferences in the visual system.

Our spatio-chromatic prefiltering mimicked the opponency of small bistratified ganglion cells and of midget cells under the assumption of a cone-type unspecific wiring of the surround. However, the exact composition of the surround of retinal receptive fields is unclear (Reid and Shapley, 1992, 2002; Lee et al., 1998, 2012; Martin et al., 2001; Buzás et al., 2006; Field et al., 2010; Crook et al., 2011; Martin et al., 2011). We determined how different surround compositions affect the distributions of chromatic preferences and the sparseness of the coding. We repeated the analysis with the same parameters but with different surround structures in the filtering stage. Besides the unspecific, mixed LMS surround we also used an unspecific mixed LM surround, a cone-type specific surround, and intermediate (mixed but biased) models for the surround (cf. **Table 1**). In addition we also used spatio-chromatic decorrelation via Zero-phase Whitening (Bell and Sejnowski, 1997). Compared to whitening, plausible retinal filtering led to more uniform distributions. Among the considered variants of retinal filtering, an equally balanced mixed surround with contributions from all cone types resulted in the smallest deviation from uniformity, but the other surround structures yielded similar values (cf. **Table 1**). Our results therefore do not provide a strong indication in favor of any specific surround organization. This suggests that in the real visual system there might be a high variation in the surround composition, which could explain why experimental evidence on the specificity of the surround has so far not been conclusive.

In addition to having more distributed preferences, the learned code for preprocessed data had all attributes one would expect from a sparse code in the cortex. The lifetime sparseness of individual components was high, but lower than in the case of unfiltered LMS data (10.29 vs. 21.40). On the other hand, the population kurtosis was drastically increased (12.67 vs. 4.86), meaning that only a small subset of all available units were active

Table 1 | Kullback-Leibler-Divergence from uniformity, mean lifetime kurtosis $\overline{K_L}$ and mean population kurtosis $\overline{K_P}$ for different surround configurations and preprocessing methods.

Surround organization	Surround for center cone of type			$D_{u(O)}$	$\overline{K_L}$	$\overline{K_P}$
	L-center	M-center	S-center			
Mixed LMS	$\frac{1}{3}(L+M+S)$			16.65	10.29	12.67
Mixed LM	$\frac{1}{2}(L+M)$			16.81	9.09	10.35
Specific	M	L	$\frac{1}{2}(M+S)$	17.70	9.25	10.80
Biased LM	$\frac{1}{2}(\frac{1}{2}(L+M)+L)$	$\frac{1}{2}(\frac{1}{2}(L+M)+M)$	$\frac{1}{2}(\frac{1}{2}(L+M)+S)$	20.23	8.86	10.04
Biased LMS	$\frac{1}{2}(\frac{1}{3}(L+M+S)+L)$	$\frac{1}{2}(\frac{1}{3}(L+M+S)+M)$	$\frac{1}{2}(\frac{1}{3}(L+M+S)+S)$	21.27	9.40	10.37
Whitening	Spatio-chromatic whitening via ZCA			32.11	9.40	4.86
None	No pre-filtering and rectification (pure L,M,S signals)			25.58	21.40	4.84

Rows 1–5 show results when surround configurations were altered while holding all other parameters constant. Additionally the results are shown for when spatio-chromatic filtering via zero-phase whitening (Bell and Sejnowski, 1997) was performed instead of center-surround filtering (row six). The last row shows the results obtained from non-preprocessed pure LMS signals as in Lee et al. (2002).

at the same time. This fits very well with the vast increase in number of neurons from LGN to visual cortex, which is paralleled in our study by the increase in dimensionality. Moreover, the code revealed by our analysis was also highly disperse, i.e., for different stimuli different subsets of units were active. This is in contrast to a compact code like PCA, where also only a few components are active all the time, but it is always the same components that take part in the coding. Such an unequal distribution of activity would seem biologically implausible because a majority of the neurons would be there without making substantial contributions to the encoding of the stimuli.

A substantial amount of ATAs (14.3%) had chromatic selectivities that corresponded to variation between light-blue and dark-yellow. Moreover, the overall distribution of color preferences also varied along one main direction in color space. Both of these axes of variation were very close to the perceptual blue-yellow axis and the line of variation of natural daylight illuminations (Mollon, 2006), which constitutes the main chromatic variation of natural scenes (Webster and Mollon, 1997) and was found in previous ICA analyses (Wachtler et al., 2001). It it also reflected in the peak of the distribution of color preferences in primary visual cortex (Wachtler et al., 2003). Our results support the conclusion that the statistics of natural scenes are an important factor in shaping the processing mechanisms of the visual system.

ACKNOWLEDGMENTS

We thank Garrett Greene and Christian Leibig for fruitful discussions and critical comments on the manuscript, Delwen Franzen for proofreading, and the reviewers for constructive comments which helped to improve the manuscript.

Supported by the Munich Graduate School of Systemic Neurosciences (GSN) and the Bernstein Center for Computational Neuroscience Munich (BMBF grant 01GQ1004A).

REFERENCES

Barlow, H. B. (1961). "Possible principles underlying the transformations of sensory messages," in *Sensory Communication*, ed W. A. Rosenblith (Cambridge, MA: MIT Press), 217–234.

Barlow, H. B. (2001). Redundancy reduction revisited. *Network* 12, 241–253.

Bell, A. J., and Sejnowski, T. J. (1997). The "independent components" of natural scenes are edge filters. *Vision Res.* 37, 3327–3338. doi: 10.1016/S0042-6989(97)00121-1

Bethge, M. (2006). Factorial coding of natural images: how effective are linear models in removing higher-order dependencies? *J. Opt. Soc. Am. A. Opt. Image Sci. Vis.* 23, 1253–1268. doi: 10.1364/JOSAA.23.001253

Boycott, B., and Wässle, H. (1999). Parallel processing in the mammalian retina. *Invest. Ophthalmol. Vis. Sci.* 40, 1313–1327.

Buzás, P., Blessing, E. M., Szmajda, B. A., and Martin, P. R. (2006). Specificity of M and L cone inputs to receptive fields in the parvocellular pathway: random wiring with functional bias. *J. Neurosci.* 26, 11148–11161. doi: 10.1523/JNEUROSCI.3237-06.2006

Caywood, M. S., Willmore, B., and Tolhurst, D. J. (2004). Independent components of color natural scenes resemble V1 neurons in their spatial and color tuning. *J. Neurophysiol.* 91, 2859–2873. doi: 10.1152/jn.00775.2003

Crook, J. D., Manookin, M. B., Packer, O. S., and Dacey, D. M. (2011). Horizontal cell feedback without cone type-selective inhibition mediates "red-green" color opponency in midget ganglion cells of the primate retina. *J. Neurosci.* 31, 1762–1772. doi: 10.1523/JNEUROSCI.4385-10.2011

Dacey, D. M. (2000). Parallel pathways for spectral coding in primate retina. *Annu. Rev. Neurosci.* 23, 743–775. doi: 10.1146/annurev.neuro.23.1.743

Derrington, A. M., Krauskopf, J., and Lennie, P. (1984). Chromatic mechanisms in lateral geniculate nucleus of macaque. *J. Physiol. (Lond.)* 357, 241–265.

Doi, E., Inui, T., Lee, T.-W., Wachtler, T., and Sejnowski, T. J. (2003). Spatio-chromatic receptive field properties derived from information-theoretic analyses of cone mosaic responses to natural scenes. *Neural Comput.* 15, 397–417. doi: 10.1162/089976603762552960

Eichhorn, J., Sinz, F., and Bethge, M. (2009). Natural image coding in V1: how much use is orientation selectivity? *PLoS Comput. Biol.* 5:e1000336. doi: 10.1371/journal.pcbi.1000336

Field, G. D., Gauthier, J. L., Sher, A., Greschner, M., Machado, T. A., Jepson, L. H., et al. (2010). Functional connectivity in the retina at the resolution of photoreceptors. *Nature* 467, 673–677. doi: 10.1038/nature09424

Foster, D. H., Amano, K., Nascimento, S. M. C., and Foster, M. J. (2006). Frequency of metamerism in natural scenes. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* 23, 2359–2372. doi: 10.1364/JOSAA.23.002359

Hoyer, P. O., and Hyvärinen, A. (2000). Independent component analysis applied to feature extraction from colour and stereo images. *Network* 11, 191–210. doi: 10.1088/0954-898X/11/3/302

Jutten, C., and Herault, J. (1991). Blind separation of sources, part I: an adaptive algorithm based on neuromimetic architecture. *Signal Process.* 24, 1–10. doi: 10.1016/0165-1684(91)90079-X

Lee, B. B. (2011). Visual pathways and psychophysical channels in the primate. *J. Physiol. (Lond.)* 589(Pt 1), 41–47. doi: 10.1113/jphysiol.2010.192658

Lee, B. B., Kremers, J., and Yeh, T. (1998). Receptive fields of primate retinal ganglion cells studied with a novel technique. *Vis. Neurosci.* 15, 161–175. doi: 10.1017/S095252389815112X

Lee, B. B., Martin, P. R., and Grünert, U. (2010). Retinal connectivity and primate vision. *Prog. Retin. Eye Res.* 29, 622–639. doi: 10.1016/j.preteyeres.2010.08.004

Lee, B. B., Shapley, R. M., Hawken, M. J., and Sun, H. (2012). Spatial distributions of cone inputs to cells of the parvocellular pathway investigated with cone-isolating gratings. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* 29, A223–A232. doi: 10.1364/JOSAA.29.00A223

Lee, T.-W., and Lewicki, M. S. (2000). "The generalized gaussian mixture model using ICA," in *Int. workshop on ICA*, (Helsinki), 239–244.

Lee, T.-W., Wachtler, T., and Sejnowski, T. J. (2002). Color opponency is an efficient representation of spectral properties in natural scenes. *Vision Res.* 42, 2095–2103. doi: 10.1016/S0042-6989(02)00122-0

Lennie, P., Krauskopf, J., and Sclar, G. (1990). Chromatic mechanisms in striate cortex of macaque. *J. Neurosci.* 10, 649–669.

Lewicki, M. S. (2002). Efficient coding of natural sounds. *Nat. Neurosci.* 5, 356–363. doi: 10.1038/nn831

Martin, P. R., Blessing, E. M., Buzás, P., Szmajda, B. A., and Forte, J. D. (2011). Transmission of colour and acuity signals by parvocellular cells in marmoset monkeys. *J. Physiol. (Lond.)* 589(Pt 11), 2795–2812. doi: 10.1113/jphysiol.2010.194076

Martin, P. R., Lee, B. B., White, A. J., Solomon, S. G., and Rüttiger, L. (2001). Chromatic sensitivity of ganglion cells in the peripheral primate retina. *Nature* 410, 933–936. doi: 10.1038/35073587

Mollon, J. D. (1989). "Tho' she kneel'd in that place where they grew..." The uses and origins of primate colour vision. *J. Exp. Biol.* 146, 21–38.

Mollon, J. D. (2006). Monge: the verriest lecture, Lyon, July 2005. *Vis. Neurosci.* 23, 297–309. doi: 10.1017/S0952523806233479

Nascimento, S. M. C., Ferreira, F. P., and Foster, D. H. (2002). Statistics of spatial cone-excitation ratios in natural scenes. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* 19, 1484–1490. doi: 10.1364/JOSAA.19.001484

Olshausen, B. A., and Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature* 381, 607–609. doi: 10.1038/381607a0

Párraga, C. A., Brelstaff, G., Troscianko, T., and Moorehead, I. R. (1998). Color and luminance information in natural scenes. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* 15, 563–569. doi: 10.1364/JOSAA.15.000563

Reid, R. C., and Shapley, R. M. (1992). Spatial structure of cone inputs to receptive fields in primate lateral geniculate nucleus. *Nature* 356, 716–718. doi: 10.1038/356716a0

Reid, R. C., and Shapley, R. M. (2002). Space and time maps of cone photoreceptor signals in macaque lateral geniculate nucleus. *J. Neurosci.* 22, 6158–6175.

Ruderman, D. L., Cronin, T. W., and Chiao, C. C. (1998). Statistics of cone responses to natural images: implications for visual coding. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* 15, 2036–2045. doi: 10.1364/JOSAA.15.002036

Schölkopf, B., and Smola, A. J. (2002). *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. Cambridge, MA: MIT Press.

Solomon, S. G., and Lennie, P. (2007). The machinery of colour vision. *Nat. Rev. Neurosci.* 8, 276–286. doi: 10.1038/nrn2094

Stockman, A., MacLeod, D. I., and Johnson, N. E. (1993). Spectral sensitivities of the human cones. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* 10, 2491–2521. doi: 10.1364/JOSAA.10.002491

Wachtler, T., Doi, E., Lee, T.-W., and Sejnowski, T. J. (2007). Cone selectivity derived from the responses of the retinal cone mosaic to natural scenes. *J. Vis.* 7, 6.

Wachtler, T., Lee, T.-W., and Sejnowski, T. J. (2001). Chromatic structure of natural scenes. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* 18, 65–77. doi: 10.1364/JOSAA.18.000065

Wachtler, T., Sejnowski, T. J., and Albright, T. D. (2003). Representation of color stimuli in awake macaque primary visual cortex. *Neuron* 37, 681–691. doi: 10.1016/S0896-6273(03)00035-7

Webster, M. A., and Mollon, J. D. (1997). Adaptation and the color statistics of natural images. *Vision Res.* 37, 3283–3298. doi: 10.1016/S0042-6989(97)00125-9

Willmore, B., and Tolhurst, D. J. (2001). Characterizing the sparseness of neural codes. *Network* 12, 255–270.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 19 June 2013; paper pending published: 19 July 2013; accepted: 04 September 2013; published online: 27 September 2013.

Citation: Kellner CJ and Wachtler T (2013) A distributed code for color in natural scenes derived from center-surround filtered cone signals. Front. Psychol. ***4****:661. doi: 10.3389/fpsyg.2013.00661*

This article was submitted to Perception Science, a section of the journal Frontiers in Psychology.

Copyright © 2013 Kellner and Wachtler. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

3

POPULATION CODING MODEL OF COLOR

All models are wrong but some are useful

— George Box, 1979

3.1 SUMMARY

Mounting evidence from physiology and theoretical work suggests that color is represented by a population code in the early visual cortex: Color selective neurons respond to a large variety of different colors and have broad, overlapping tuning curves but each responds maximally to a specific color. These preferences are widely distributed around the hue circle in the colorspace.

The response of neurons to stimuli inside their receptive field can be influenced by stimuli outside their receptive field. Such contextual modulation of neurons in early visual cortex is thought to underlie perceptual context effects, such as the tilt effect (which alters the perception of orientation), and human color induction effects. In the latter, a chromatic surround influences the perception of a test stimulus. The resulting perceptual hue shift is systematic, depends on the distance between stimulus and surround hue in colorspace, and is always away from the surround. The magnitude and shape of the hue induction curves depend on the surround hue.

Here, we model a population code for color to predict human induction effects, as measured by a psychophysics experiment. The population code consisted of color selective units with bell shaped tuning curves and uniformly distributed tuning preferences. Contextual modulation was modeled as inhibition of the unit's activation. The presence of surround modulation altered the pattern of activation in the ensemble of neurons and changed the hue encoded in the population code in a way that depended on the specific modulation. The shape the modulation

that was modeled followed a Gaussian profile with two parameters: the modulation amplitude and modulation width. Carefully selected pairs of modulation amplitude and width, fitted to psychophysical measurements, could successfully predict human hue induction curves.

3.2 CONTRIBUTIONS

The contributions of the authors Christian Johannes Kellner (CJK) and Thomas Wachtler (TW) are as follows: CJK and TW designed the study. CJK implemented and performed the simulation. CJK and TW wrote the manuscript. Earlier versions of the model results were presented at the *Society for Neurosciences Meeting 2014* in Washington, DC and the International Colour Vision Society Symposium 2015 in Sendai, Japan.

Population coding model of color in visual cortex

Christian Johannes Kellner[1,2,3]*, Thomas Wachtler[1,3]

Abstract
The perceived color of an object depends not only on the spectral composition of the light reflected from its surface, but also on the visual context such as illumination and surrounding colors. Such contextual modulations are thought to underlie perceptual phenomena such as color constancy. A possible neuronal basis may be lateral interactions that alter the activity of color selective neurons in the early visual cortex. Here, we present a model of cortical color processing that predicts color shifts induced by chromatic surrounds. The model assumes that stimulus hue is encoded by a population of neurons with Gaussian tuning curves and color preferences distributed in color space. This is based on empirical findings of distributed color preferences in primary visual cortex. The influence of the surround hue on the stimulus was implemented via inhibition of the units' activation in the population code. The strength of this inhibition depended on the distance between the stimulus and the surround hue in colorspace. This contextual modulation by the surround hue introduced a bias in the population response that, when decoded, revealed systematic shifts in the encoded stimulus hue. The strength and shape of the resulting simulated induction curve strongly depended on the modulation parameters, such as width and amplitude. The results indicate that important computations in color vision that lead to perceptual hue shifts can be realized using simple neural mechanisms when color is represented by a distributed code.

Keywords
ColorVision — ContextualModulation — PopulationCoding

[1] *Department Biologie II, Ludwig-Maximilians-Universität München, Munich, Germany*
[2] *Graduate School of Systemic Neurosciences, Ludwig-Maximilians-Universität München, Munich, Germany*
[3] *Bernstein Center for Computational Neuroscience, Munich, Germany*
***Corresponding author**: kellner@bio.lmu.de

Introduction

Population codes, where an ensemble of neurons encode a feature together, are widespread and can be found in various modalities of many animals[1]. Prominent examples supported by empirical evidence include the encoding of motor commands in the motor cortex[2], place cells and grid cells in the hippocampus, the auditory system[3] and various features in the visual modality such as orientation[4], direction of motion[5] and color[6]. Theoretical treatment of population codes in computational neuroscience ranges from the analysis of coding properties and information processing of such ensemble codes to models predicting behavior and perception[7, 8]. A famous and well studied example is the tilt effect, where the perceived orientation of a test patch with a grating is influenced by another surrounding grating. Depending on the angle of the orientation between the gratings, the perceived orientation of the center patch is shifted away or attracted towards the surround orientation[9]. Various models based on population codes have been put forward to predict the tilt effect[10]. Such contextual effects can also be found in other domains. Recently, such a contextual effect was quantified in color vision[11, 12]: When a chromatic stimulus was presented on a colored surround, the perceived hue of the stimulus was altered by the surround. These changes in perceived hue differed in strength for different surround hues but consistently shifted the perceived hue away from the surround hue when hue was expressed in cone-opponent colorspace. Additionally, electrophysiological evidence shows that neurons in early visual cortex, which are involved in color appearance, change their firing in the presence of a chromatic surround[6].

Here we present a model for color that, inspired by the theoretical work on the orientation tilt effect, uses contextual modulations of activity in a population code of color selective neurons to predict human hue induction effects.

Methods

Colorspace

All chromaticities are defined in a two dimensional cone-opponent colorspace[13], where the two coordinate axes correspond to the two chromatic pathways, L-M and S. In this colorspace the angle relative to the L-M axis ϕ determines the hue, and the distance from the center point indicates the contrast. All stimuli and surround hues are therefore assumed to be isoluminant and of the same saturation.

Population code

The population code model consists of color selective units, whose tuning, i.e. their response to different stimuli, is modeled with a circular Gaussian with three parameters. These are, a) the preferred direction of the unit at a hue angle of θ

in colorspace, b) the tuning width ω and c) its maximum response r_{max}. The basis of the circular Gaussian is a Gaussian distribution with mean μ, variance σ and a scaling factor α as:

$$G(x|\mu,\sigma,A) = A * e^{\frac{-(x-\mu)^2}{2\sigma^2}} \tag{1}$$

In order to approximate a circular Gaussian $C(x)$ that wraps around at the "edges" of the colorspace, the preferred tuning is offset by ±360° and then summed up:

$$\begin{aligned} C(x|\mu,\sigma,\alpha) =& G(x|\mu,\sigma,\alpha)+ \\ & G(x|\mu+360^\circ,\sigma,\alpha)+ \\ & G(x|\mu-360^\circ,\sigma,\alpha) \end{aligned} \tag{2}$$

For tuning curves, the maximal response r_{max} of the Gaussian is chosen so that it has unit area. This is done with a scaling factor α as follows:

$$\alpha = \frac{1}{\sigma\sqrt{2\pi}}. \tag{3}$$

The activation $u(x|\mu,\sigma)$ of an individual unit for a specific hue θ is given by:

$$u(\theta|\phi,\sigma) = C(\theta|\phi,\sigma,\alpha). \tag{4}$$

A total of 1440 units were modeled in all simulations shown in this paper. Their preferred tuning directions were uniformly distributed over the whole hue circle. The preferred angle ϕ_i of the unit i is therefore $\phi_i = \frac{360^\circ}{1440} i$. The *population activation* for a stimulus θ is the vector of the individual unit activations:

$$\vec{u}(\theta) = [u_1(\theta|\phi_1),...,u_i(\theta|\phi_i),...]^\mathsf{T}. \tag{5}$$

Surround Modulation

We modeled the modulation of a chromatic surround as multiplicative inhibition with a circular Gaussian profile. With a modulation amplitude a and modulation width w the modulated activity m_i of the unit i with a surround hue of angle θ_s is then given by:

$$m_i = u_i[1 - aC(\phi_i|\theta_s,w,1)]. \tag{6}$$

Or expressed in terms of the population activation:

$$\vec{m}(\theta|\theta_s) = [m_1,...,m_i,...]^\mathsf{T}. \tag{7}$$

Decoding

As a result of the surround modulation in the population code, the activity of the units might have changed in such a way that the encoded hue is different from that initially encoded. This change in hue is the *induced hue change* by the surround. An appropriate decoding scheme had to be chosen to read out the hue angle represented in the population activity. For all results in this paper the *maximum likelihood decoder (MLD)* was used. But for comparison several other decoders were used as well (see also Discussion).

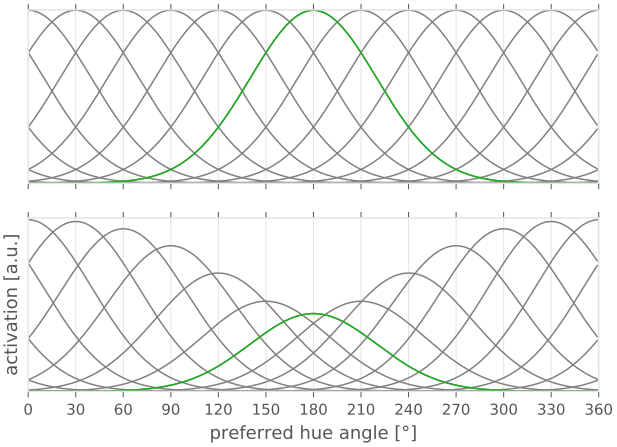


Figure 1. Contextual modulation of population code units. TOP| Tuning of a small subset of twelve color selective model units, i.e. their response as a function of stimulus hue. The preferred directions, spaced 30° apart, is indicated by light-gray vertical lines. An example unit with a preferred hue at 180° is highlighted in green. BOTTOM| Response of the same model units in the presence of a surround modulation. The activation of each unit is uniformly reduced. The reduction follows a Gaussian profile centered at the surround hue and is therefore strongest when surround and stimulus hue are the same.

The basic assumption of the **maximum likelihood decoder (MLD)** decoder is that downstream neurons do not "know" if a modulating surround is present or not. From this, it follows that patterns that are similar to each other, regardless of whether they result from modulated activity or not, should be decoded similarly, thus leading to the same percept. Specifically, a stimulus s_g on a neutral gray surround leading to activation $\vec{u}_g$ and another stimulus s_c on a colored surround leading to the activation profile $\vec{m}_m$ should lead to similar decoded hues if $\vec{m}_m$ is very similar to $\vec{u}_g$. Or, in mathematical terms, we want to maximize the conditional probability of the activation $\vec{m}$ given the angle θ: $h = \arg\max_\theta P(\vec{m}|\theta)$, where h is then the estimated angle [8]. The conditional probability is also called the likelihood of θ, and hence the decoder is called the *maximum likelihood decoder*. The implementation of such a decoder was done via lookup table with a simple vector similarity measure. This table is first created by uniformly sampling the entire hue circle with N samples, where the hue angle h_i corresponding to the entry e_i in the lookup table is given by $h_i = \frac{360^\circ}{N}$. The entry e_i in the lookup table is then created by obtaining the population activity without surround modulation (i.e. assuming a gray surround hue). The entry most similar to the given modulated population activity is selected. Its corresponding angle h is the decoded hue. Similarity is given by the Euclidean distance (L^2-norm) between the two vectors. The resolution limit is $360^\circ/N$. For all

simulations in this study N was 3600, i.e. the resolution limit was 0.1°.

The simplest alternative decoding scheme is the **maximum fire-rate** decoder. Here, the decoded hue is given by the preferred angle ϕ of the unit with the maximum activation. Such a decoder could, for example, be realized by a winner-takes-all network. The resolution limit of this decoder is the spacing between the individual units and therefore depends on the number of units in the population code. For a population code with N units of uniformly distributed preferred directions, the limit is $360°/N$. In our case with $N = 1440$ the resolution limit was 0.25°.

A standard decoder that is widely used in the field is the **population vector** [2]. The decoded hue is acquired by the vector sum of the preferred direction of each unit weighted by their respective activation:

$$h = \arctan 2(\sum_i m_i * \sin(\phi_i^r), \sum_i m_i * \cos(\phi_i^r)) \qquad (8)$$

Where $\phi_i^r = \phi_i \frac{\pi}{180°}$ is the preferred direction of unit i converted to radian, and m_i its activation after modulation.

Another decoding scheme involves fitting the activation profile with a circular Gaussian $C(x|\mu,\sigma,A)$ (**Gaussfit**). Free parameters to fit are μ, σ and A. The sum of squared differences between the population activity and the circular Gaussian was minimized using SciPy's *scipy.optimize.leastsq*. The decoded hue angle is the parameter μ after fitting. The parameter μ was initialized with the preferred angle of the unit with the maximum activation. The paramter σ was set to the tuning width of that unit and A was set to 1.0. Here, the resolution limit is not directly constrained but also indirectly depends on the number of units, since this influences the quality of the fit.

Model Parameter and Empirical Data

The findings of Klauke et al.[11] served as the empirical basis of our study. In their psychophysics study hue induction effects were measured for eight different, equally spaced surrounds and 16 different, equally spaced test stimuli. The latter were given in hue angles relative and centered to the surround, therefore splitting the colorspace in two halves, i.e. the eight positive and eight negative stimulus hue angles (-180° to 0° and 0° to +180°, NB: -180° and +180° are the same absolute hue angle). For direct comparison, the same surround and test stimulus angles were used in our simulations.

The three free parameters of our population code model are: the tuning width of the population units σ, the modulation amplitude and the modulation width. A parameter scan was performed where σ and modulation width w were varied from 10° to 100° in steps of 2°; the modulation amplitude a was varied between 0.1 and 0.9 in steps of 0.05. For each combination of σ, a, m, we simulated the population code with surround modulations of the eight surround hues.

To quantify the differences between simulated hue changes and experimental data, we sampled hue induction shifts at the same locations as the empirical data. At each location we computed the squared difference and multiplied the result with the absolute experimental hue shift. This gives more weight to locations with high induction effects. Finally, for both the positive and negative halves (see above) we summed up the weighted squared difference to obtain one single value indicating the difference between model and empirical data.

Parameter selection was done by finding the combination of parameters that gave the minimum difference between model and empirical data. First, the tuning width, σ, was estimated from the simulations and then fixed: For each parameter set consisting of surround hue θ_s, modulation amplitude a, modulation width w, and σ, we computed the average differences of positive and negative halves. For each combination of background and σ we selected the pair of modulation parameters a, w giving the minimal differences. Then for each σ the mean difference was computed by averaging over all backgrounds. The sigma giving the minimal mean difference was subsequently used.

To asses how well the population code model can predict the psychophysics data we calculated the square-root of the mean of squared differences between model and experimental data for each surround separately (root-mean-square per surround, RMS_s). The mean of the surround specific root mean square values was then used as a single metric of the overall accuracy of the model.

Software

Software packages used for the data analysis include Python (2.7.11), NumPy[14] (1.11.0), SciPy[15] (0.17.1) and Pandas[16] (0.14.1). Plots were generated with Matplotlib[17] (1.5.1).

Results

In the presence of a colored surround the response of individual model units is modulated as a function of distance between the unit's preferred direction and the surround hue. Figure 1 (TOP) shows the activation of a subset of units to test stimuli spanning the whole hue circle, i.e. their tuning. The BOTTOM panel of the same figure depicts the responses of the identical units to the same test stimuli, but now under the influence of the surround modulation. The activity of each unit is uniformly reduced up to a maximum in the case where the unit's preferred direction is equal to the surround hue.

At the population level, this also leads to a reduction in overall activity. Figure 2 depicts the population activity before (green) and after (red) the surround modulation. In the case where surround modulation is equal to the test stimulus, the activation profile of the population units is uniformly reduced. The encoded hue does not change. If the surround hue is slightly different from the test stimulus, the modulation affects the flank of the activation profile closest to it while the other flank is less or not at all affected. This changes the shape of the profile so that it becomes asymmetrical. More importantly, the peak of the population activity is also shifted *away* from the surround hue. All this leads to a change in the hue that

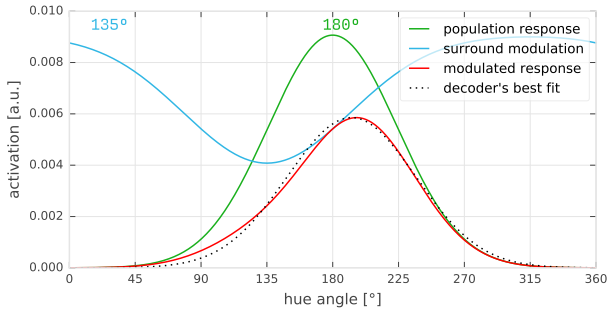


Figure 2. Population activity. The activation profile of the population in the absence of a surround modulation (i.e. gray surround) for the test stimulus of 180° is shown in green. Since all tuning curves had the same Gaussian tuning the activation profile is itself Gaussian. The light blue curve depicts the reduction in activity at each location for a surround at 135° (normalized to the maximum activity). The resulting modulated population activity is shown in red. It is clear that the peak of the activation profile is shifted away from the surround location; additionally the activation profile becomes asymmetric. The dotted orange curve is the most similar activation profile found in the lookup table of the MLD decoder (after all profiles in the lookup table had been normalized to the maximum firing of the modulated population).

is encoded in the population activity, which depends on the distance in colorspace between surround and stimulus hue.

The difference in the encoded hue between modulated and unmodulated population activity, i.e. the hue change induced by the surround, is similar to the perceptual hue shifts induced by chromatic surrounds in asymmetric color matching experiments[11, 12]. In these experiments, induced shifts depended on the surround chromaticity. The strength of the induced hue in the model depends not only on the distance between surround and stimulus hue, but also on the modulation depth and width (and also on the decoder used to read out the encoded hue, see Discussion for details). While the distance between stimulus and surround hue is responsible for the stereotypical sinusoidal-like shape, the other two parameters determine the peak and the slope of the curve. The different shapes for different surrounds found in psychophysics studies can be reproduced with these two parameters.

Figure 3 shows such hue induction curves from one study[11] for eight different surrounds as a function of the stimulus hue, together with simulation results from the population code model. The model parameters (modulation amplitude and width) were fitted separately for each background and for both halves of the hue induction curves. In general, the overall shape of the individual half curves can be reproduced adequately. Peak location, initial slope up to the peak, and the declining slope from the peak to zero are faithfully reproduced in most cases. However, fits were not optimal in a few locations, most prominently the positive half curve for the surround of 0° hue angle. The peak location as well as the downward slope are off by several degrees. A possible explanation is that while in the empirical data the peak of the induced hues is very high (around 20°) and located relatively far away from the surround hue (around 65°), in the model the higher the peak is, the closer it is to the surround hue. The average prediction accuracy, as measured by the root mean square difference of simulation and data, was 2.54°±0.65° for all surrounds combined. The best prediction accuracy was achieved for the 90° surround with a RMS of 1.59°, while the 45° surround had the worst fit (RMS of 3.84°). The reason for this is that although the slope and peak height were reproduced quite well, the psychophysics curve exhibits a positive offset, i.e. it does not run through the origin, that cannot be reproduced by the model. The RMS_s values for the individual surrounds, with positive and negative halves combined are: 0°: 3.14°, 45°: 3.84°, 90°: 1.59°, 135°: 2.12°, 180°: 2.14°, 225°: 2.68°, 270°: 2.39°, 315°: 2.45°.

The modulation parameters for the fits in Figure 3 are shown in Figure 4. Modulation amplitude (TOP) and width (BOTTOM) are plotted as a function of the surround hue for both halves separately. The maximum and minimal modulation amplitudes were 0.85 and 0.45, and the mean value was 0.67±0.1. The modulation width spanned the full parameter scan range, from the minimal modulation width of 10° (surround hue 315°, negative half) up to 98° (e.g. surround hue 315°, positive half), with a mean of 62.75°±31.8. If modulation amplitude is high and width is narrow, e.g. positive half of the 45° surround, the resulting hue induction curve has a high peak close to the origin, resulting in a steep slope. The ratio of modulation amplitude and width $r = \frac{w}{a}$ for 45°+ is $r_{45°+} = 23.52$. In such a situation, the value of the peak itself seems not to depend as much on the ratio, since in the case of 135° negative half, the ratio is similar, $r_{135°-} = 24.62$. Rather it depends more on the absolute values, which are both smaller in the case of 135°-: $(w,a)_{45°+} = (0.85, 20°)$ vs $(w,a)_{135°-} = (0.75, 10°)$. If both parameters have small values, the resulting curve is shallow, with a small peak around 45° away from the origin.

Additionally, the tuning width ω of the model units was a free parameter. In order to not make the model too complex and constrain the influence of the surround to one mechanism, i.e. the modulation of the activity of the unit, we set the parameter ω to be the same for all backgrounds. It was selected from a range of 10° to 100° in a way such that it yielded the best average fit for all eight surrounds for all pairs of modulation width and amplitude. This resulted in a tuning width of 40°. As a comparison, if ω was allowed to vary freely and be different for each surround and induction half curve, the average root mean square prediction error decreased by 0.19° to 2.35°±0.65°, ranging from 1.43° to 3.78°. The values for the various surrounds are: 0°: 2.70°, 45°: 3.78°, 90°: 1.43°, 135°: 2.11°, 180°: 1.79°, 225°: 2.45°, 270°: 2.31°, 315°: 2.27° (positive and negative halves combined). The tuning width ω itself was in the range of 18° to 94°.

If only one set of parameters was fitted for each of the surrounds, the average RMS value increased to 3.11°±0.64°. Per surround RMS_s values are, 0°: 3.83°, 45°: 3.60°, 90°:

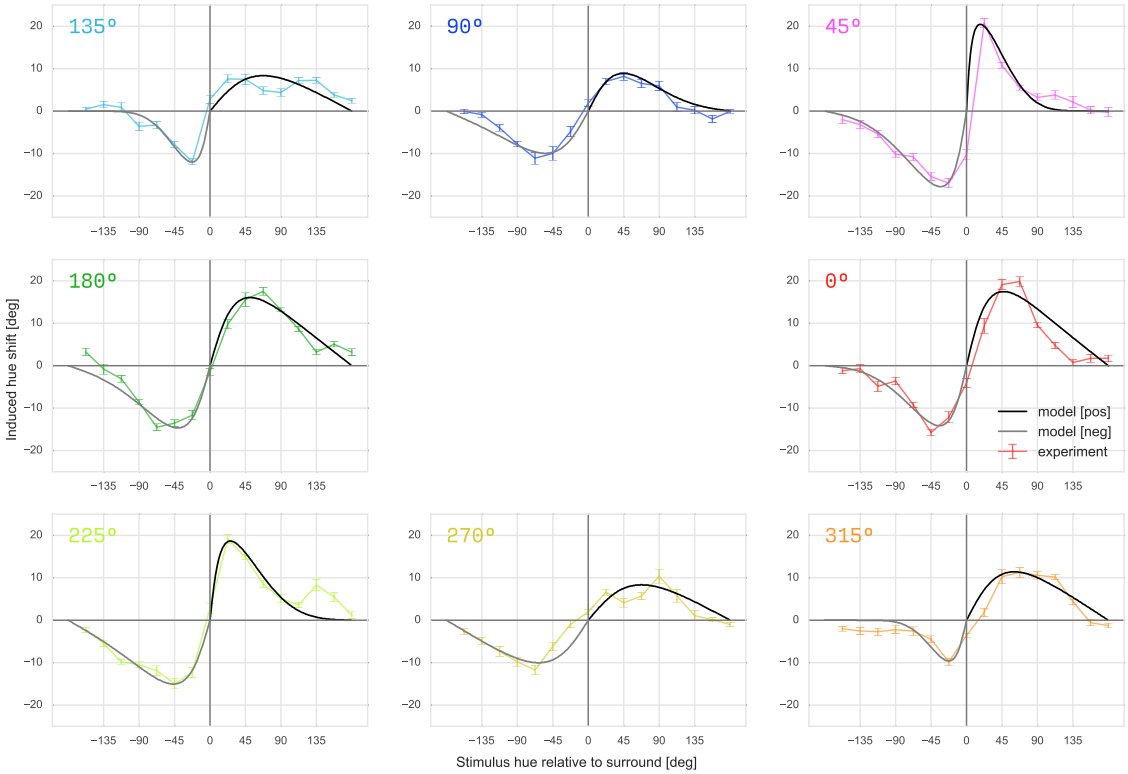


Figure 3. Comparison of simulated hue changes to empirical data for eight different surrounds. Each panel shows the hue induction changes for a colored surround as a function of test stimulus hue. Data from the empirical study[11], plotted in different colors, are averages from five different subjects; error bars indicate the standard error. Modulation width and amplitude of the model were fitted separately for positive and negative halves and the resulting simulated hue shifts are shown in black for positive and gray for negative halves.

1.81°, 135°: 3.21°, 180°: 2.95°, 225°: 3.05°, 270°: 2.63°, 315°: 3.83° (minimum: 1.81° and maximum 3.83°).

Discussion

In this study we describe a population code model of color selective units whose activity is modulated by a chromatic surround, and use it to predict human color induction effects. If modulation width and depth are varied, all shapes of the induction shift curves can be fitted.

Population code configuration

The fundamental assumption in our model is that color in visual cortex is represented by a population of color selective units. This is based on empirical physiological data obtained in primates[6, 18–21], imaging data from humans[22], and theoretical work based on natural scene statistics[23, 24]. All these studies found that peak color preferences of neurons were widely distributed and many cells responded to the same hue albeit with different rates. All this is in theory not a problem for the model but leads to an explosion of parameters. Ultimately, many different configurations of the population code could lead to the same population activity. In contrast to our idealistic model of uniformly distributed peaks and units with symmetric, bell-shaped tuning curves, empirical data suggests that the distribution is neither uniform, nor are tuning curves so well behaved. Instead, clusters have been identified[19] and tuning widths of neurons seem to be distributed over a broad range (cf. e.g. [6]). Some aspects of a more biologically plausible model might be necessary to explain certain observations, such as the zero-crossing shift of hue induction curves. In psychophysical experiments the surround hue and the center stimulus had slightly different saturation[11, 12], which might be the cause for this shift. Thus the model needs to be extended to include saturation tunings.

The tuning width ω of the model units in the population code was selected during the fitting process. From ω values in the range of 10° to 100° we selected the one that had the highest prediction accuracy, as measured by the root mean square deviation from the empirical data, averaged over all surrounds. The full width at half maximum (FWHM) of the tuning curves can be expressed in terms of the tuning width ω: $\mathrm{FWHM} = 2\sqrt{2\ln 2} \approx 2.3548\omega$. The estimated tuning width in our model was 40°, which corresponds to a FWHM of approximetly 94°. The distribution of tuning widths of

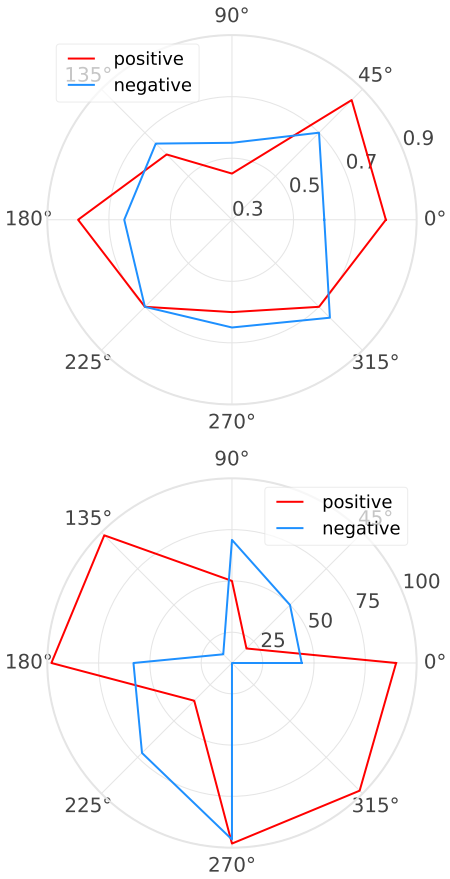


Figure 4. Modulation parameters. Modulation amplitude (TOP) and width (BOTTOM) plotted as a function of surround hue for the positive (red) and negative (blue) induction half curves shown in Figure 3.

neurons that respond to color stimuli measured in macaque monkeys is broad, and ranges from 30° to 120° FWHM but also shows a clear peak in the histogram around 90° [6], and is therefore in good agreement with our data.

Mechanisms of contextual modulation

Contextual modulations in our population code simulation were implemented by a uniform reduction of activity that depended on the angular hue distance of stimulus and surround. We did not simulate any mechanism between the populations of stimulus and surround neurons, since we were interested in whether the different shapes measured could theoretically be reproduced by modulating the population activity, independently of the specific mechanisms involved. Nevertheless, this of course remains an important question. Since there seems to be no filling in for color directly in primary visual cortex[25] and hue inductions are also size dependent[12], more than one mechanism will most likely be involved. Horizontal lateral interactions[26] at the color edge combined with feedback from higher stages of color processing is one likely solution.

A possible alternative starting point that could en passant also explain the anisotropic tuning widths and peak densities of neurons is that the surround dependent differences in human hue induction curves result from the colorspace used in the experiments. The colorspace is based on the three opponent mechanisms found in sub-cortical pathways and there is no straightforward connection between perception and these mechanisms. The unique hues do not coincide with the cardinal axes of the colorspace and do not even fall on a line. It is also not perceptually uniform, which means that in some areas the hue angle can undergo big changes with only a small perceptual effect. A transformation of the colorspace that leads to the same (or at least very similar) psychophysical hue induction curves could be used to transform the tuning curves and density of model units so that a uniform modulation in the resulting anisotropic population code would lead to surround specific induction curves.

Choice of decoder

The choice of the decoder used to read out the hue encoded in the population code can have a significant effect. Figure 5 (TOP) shows hue induction curves for a single population code undergoing a surround modulation and four different decoders: The MLD decoder used for all results in this paper, the Gaussfit decoding scheme, the maximum fire-rate decoder and the population vector. The Gaussfit and MLD decoders generated quite similar results, but the maximum fire-rate and the population vector showed drastic changes in both the height and location of the maximum hue shift and the general shape of the induction curves. The population vector is used in many studies that model population codes, making this decoding scheme the de-facto standard. The reason we chose the MLD decoder over any other was that it makes only one assumption about how downstream neurons process the information encoded in the population code: they are not aware that contextual modulation has happened. Under such an assumption, similar population activity should lead to similar decoded hues and thus, similar percepts. Additionally, the MLD decoder is also robust and works as expected with anisotropic population codes where the tuning curve widths are different for different units and the density of the units' preferred stimuli is non-uniformly distributed. In such a case, the maximum fire-rate and population vector decoders both show a bias that leads to hue changes even for a neutral, gray background. Figure 5 (BOTTOM) shows the response of the various decoders for an anisotropic population code with a *sine* modulated tuning density under a neutral surround condition. Both maximum fire-rate and population vector show a bias that reflects the *sine* modulation.

Conclusions

Population coding of hue with inhibitory surround modulation leads to shifts in the encoded color similar to induction

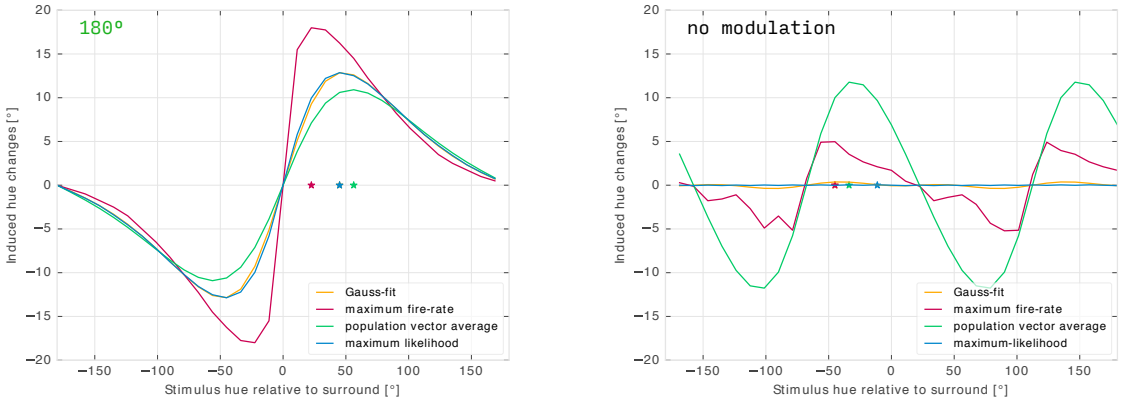


Figure 5. Comparison of different population code decoders. Both panels show hue induction curves obtained via four different population code decoders: Maximum fire-rate, Gaussfit, population vector and maximum likelihood (MLD).
TOP| A population code of units with identically shaped tuning curves and color preferences uniformly distributed across the hue circle was simulated with a surround modulation at hue angle 180°. The activation profile was read out via each decoder. Shape and peak vary drastically between the decoders with only the Gaussfit and MLD decoder yielding similar results.
BOTTOM| A population code with units of varying tuning width and non-uniformly distributed color preferences. No surround modulation was simulated, therefore the hue induction curves depict the decoder bias.

effects observed in human perception. The magnitudes of hue shifts depend on modulation amplitude and modulation width, providing an explanation for the observed induction strength variation with surround hue.

Funding

Bundesministerium für Bildung und Forschung (01GQ1004A); Munich Graduate School of Systemic Neurosciences (GSN).

References

1. Pouget, A., Dayan, P. & Zemel, R. Information processing with population codes. *Nat. Rev. Neurosci.* **1,** 125–132 (2000).
2. Georgopoulos, A. P., Schwartz, A. B. & Kettner, R. E. Neuronal population coding of movement direction. *Science* **233,** 1416–1419 (1986).
3. Grothe, B. New roles for synaptic inhibition in sound localization. *Nat. Rev. Neurosci.* **4,** 540–550 (2003).
4. Hubel, D. H. & Wiesel, T. N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol. (Lond.)* **160,** 106–154 (1962).
5. Maunsell, J. H. & Van Essen, D. C. Functional properties of neurons in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed, and orientation. *J. Neurophysiol.* **49,** 1127–1147 (1983).
6. Wachtler, T., Sejnowski, T. J. & Albright, T. D. Representation of color stimuli in awake macaque primary visual cortex. *Neuron* **37,** 681–691 (2003).
7. Averbeck, B. B., Latham, P. E. & Pouget, A. Neural correlations, population coding and computation. *Nat. Rev. Neurosci.* **7,** 358–366 (2006).
8. Pouget, A. & Latham, P. E. in *Bayesian Brain* 893–897 (MIT Press, Cambridge, MA, 2007).
9. O'Toole, B. & Wenderoth, P. The tilt illusion: repulsion and attraction effects in the oblique meridian. *Vis. Res.* **17,** 367–374 (1977).
10. Schwartz, O., Sejnowski, T. J. & Dayan, P. Perceptual organization in the tilt illusion. *J. Vis.* **9,** 19.1–20 (2009).
11. Klauke, S. & Wachtler, T. "Tilt" in color space: Hue changes induced by chromatic surrounds. *J. Vis.* **15,** 17 (2015).
12. Kellner, C. J. & Wachtler, T. Stimulus size dependence of hue changes induced by chromatic surrounds. *J. Opt. Soc. Am. A* **33,** A267–A272 (2016).
13. MacLeod, D. I. & Boynton, R. M. Chromaticity diagram showing cone excitation by stimuli of equal luminance. *J. Opt. Soc. Am. A* **69,** 1183–1186 (1979).
14. Van der Walt, S., Colbert, S. & Varoquaux, G. The NumPy Array: A Structure for Efficient Numerical Computation. *Computing in Science Engineering* **13,** 22–30. ISSN: 1521-9615 (2011).
15. Jones, E., Oliphant, T., Peterson, P., *et al. SciPy: Open source scientific tools for Python* [Online; accessed 2015-08-24]. <http://www.scipy.org/> (2001–).
16. McKinney, W. *Data Structures for Statistical Computing in Python* in *Proceedings of the 9th Python in Science Conference* (eds van der Walt, S. & Millman, J.) (2010), 51–56.

17. Hunter, J. Matplotlib: A 2D Graphics Environment. *Computing in Science Engineering* **9,** 90–95. ISSN: 1521-9615 (2007).

18. Thorell, L. G., De Valois, R. L. & Albrecht, D. G. Spatial mapping of monkey V1 cells with pure color and luminance stimuli. *Vis. Res.* **24,** 751–769 (1984).

19. Lennie, P., Krauskopf, J. & Sclar, G. Chromatic mechanisms in striate cortex of macaque. *J. Neurosci.* **10,** 649–669 (1990).

20. De Valois, R. L., Cottaris, N. P. & Elfar, S. D. Some transformations of color information from lateral geniculate nucleus to striate cortex. *Proc. Natl. Acad. Sci. U.S.A.* **97,** 4997–5002 (2000).

21. Hanazawa, A., Komatsu, H. & Murakami, I. Neural selectivity for hue and saturation of colour in the primary visual cortex of the monkey. *Eur. J. Neurosci.* **12,** 1753–1763 (2000).

22. Kuriki, I., Sun, P., Ueno, K., Tanaka, K. & Cheng, K. Hue Selectivity in Human Visual Cortex Revealed by Functional Magnetic Resonance Imaging. *Cereb. Cortex* **25,** 4869–4884 (2015).

23. Caywood, M. S., Willmore, B. & Tolhurst, D. J. Independent components of color natural scenes resemble V1 neurons in their spatial and color tuning. *J. Neurophysiol.* **91,** 2859–2873 (2004).

24. Kellner, C. J. & Wachtler, T. A distributed code for color in natural scenes derived from center-surround filtered cone signals. *Front. Psychol.* **4** (2013).

25. Zweig, S., Zurawel, G., Shapley, R. & Slovin, H. Representation of Color Surfaces in V1: Edge Enhancement and Unfilled Holes. *J. Neurosci.* **35,** 12103–12115 (2015).

26. Stettler, D. D., Das, A., Bennett, J. & Gilbert, C. D. Lateral connectivity and contextual interactions in macaque primary visual cortex. *Neuron* **36,** 739–750 (2002).

4

HUE INDUCTION IN HUMANS IS STIMULUS SIZE DEPENDENT

Daß es mir – oder Allen – so scheint,
daraus folgt nicht, daß es so ist.

— Wittgenstein, UG

4.1 SUMMARY

The perception of colors is influenced by the presence of a chromatic surround. This effect was recently quantified in a psychophysics experiment. When hue is expressed as an angle in cone-opponent colorspace, the induced hue shift was always away from the surround hue and its size depended on the distance between the two. Although qualitatively the effect was the same for different surround hues, their exact shape — including the magnitude and location of the peaks — varied with different surrounds. One open question that could guide future models is whether this effect is size dependent, if at all present for large stimuli.

Here, we repeated the initial psychophysics experiment that was done with stimuli of 2° visual angle, but now with two additional sizes: a smaller one of 0.5° visual angle and a larger one of 8° visual angle. Our results show that, for the majority of surround hues, smaller stimuli lead to larger hue induction effects, and larger stimuli result in smaller induced hue shifts. The exception were surrounds along the S-cone cardinal axis of the colorspace. The results suggest that at least two different mechanisms are involved: The generation of the actual hue shift, most likely by spatial lateral interactions in primary visual cortex, in addition to a *filling in* that is mediated by higher stages and causes the size dependence observed in the data.

The contributions of the authors Christian Johannes Kellner (CJK) and Thomas Wachtler (TW) are as follows: CJK and TW designed the study. CJK implemented and supervised the experiment, and also did the data analysis. CJK and TW discussed the data, results and wrote the manuscript. First results were presented at the *International Colour Vision Society Symposium 2015* in Sendai, Japan. The manuscript was accepted as publication in the *Journal of the Optics Society of America A* as:

Kellner, C. J. & Wachtler, T. (2016).Stimulus size dependence of hue changes induced by chromatic surrounds. *J. Opt. Soc. Am. A*. *33*(3), A267–A272. doi: 10.1364/JOSAA.33.00A267.

Journal of the Optical Society of America A
OPTICS, IMAGE SCIENCE, AND VISION

Stimulus size dependence of hue changes induced by chromatic surrounds

CHRISTIAN JOHANNES KELLNER[1,2,3,*] AND THOMAS WACHTLER[1,3]

[1]Department Biologie II, Ludwig-Maximilians-Universität München, Munich, Germany
[2]Graduate School of Systemic Neurosciences, Ludwig-Maximilians-Universität München, Munich, Germany
[3]Bernstein Center for Computational Neuroscience, Munich, Germany
*Corresponding author: kellner@bio.lmu.de

Received 8 October 2015; revised 12 December 2015; accepted 16 January 2016; posted 19 January 2016 (Doc. ID 251587); published 26 February 2016

A chromatic surround induces a change in the perceived hue of a stimulus. This shift in hue depends on the chromatic difference between the stimulus and the surround. We investigated how chromatic induction varies with stimulus size and whether the size dependence depends on the surround hue. Subjects performed asymmetric matching of color stimuli with different sizes in surrounds of different chromaticities. Generally, induced hue shifts decreased with increasing stimulus size. This decrease was quantitatively different for different surround hues. However, when size effects were normalized to an overall induction strength, the chromatic specificity was largely reduced. The separability of inducer chromaticity and stimulus size suggests that these effects are mediated by different neural mechanisms. © 2016 Optical Society of America

OCIS codes: (330.1720) Color vision; (330.5510) Psychophysics.

http://dx.doi.org/10.1364/JOSAA.33.00A267

1. INTRODUCTION

One of the most striking phenomena in the domain of color vision is how the perception of the color of an object depends not only on its own reflective properties but also on the visual context it is presented in. Prominent contextual features that influence the appearance of color are, among others, the illumination of the scene, the brightness of the background, and, notably, other colors surrounding the object. For example, a chromatic context changes the appearance of a neutral stimulus such that a gray stimulus patch looks yellowish when presented on a blue background and bluish on a yellow background. Additionally, various object dimensions influence the strength of chromatic induction, such as distance to the surround, spatial frequency, shape, and size [1,2].

The influence of chromatic surrounds on the perceived hue of color stimuli has recently been quantified by Klauke and Wachtler [3]. In line with previous studies on chromatic induction, it was shown that a chromatic surround systematically induced hue changes that shifted the perceived hue away from the surround hue, thus enhancing the perceptual color contrast [4–7].

Such color induction effects are thought to play an essential role in color constancy [8]. Evidence suggests that the underlying neural mechanisms include lateral interactions in the early visual system [2,9]. Chromatic tuning of neurons in V1 is influenced by a chromatic surround in a way that is consistent with the perceptual color induction effect [10]. As a result, various computational models for color constancy rely on local contrast that is computed at the border of surfaces. Since receptive fields in the early visual system are of limited size, this entails the question of the influence of stimulus size on the chromatic induction. In the case of achromatic stimuli, Shi *et al.* [11] recently investigated the effect of size on simultaneous contrast. However, for color no quantitative data exists yet. Here we examined the influence of stimulus size on hue shifts induced in chromatic test patches presented in colored surrounds.

2. METHODS

A. Observers

Five subjects, four males and one female, age range from 27 to 50 years, took part in the study. All observers had normal color vision as assessed by the Farnsworth–Munsell 100 Hue test and Ishihara plates and had normal or corrected to normal visual acuity. Two of the subjects were the authors, while the other subjects were naïve with respect to the purpose of the study and gave informed consent to participate in the study.

B. Apparatus and Stimuli

Stimuli were presented on a 21" Sony GDM F520 CRT screen connected via an ATI Radeon HD 4200 graphics card. Screen resolution was 1280 × 1024 pixels at a refresh rate of 85 Hz. The display was calibrated using a PhotoResearch (Chatsworth,

California) PR-655 spectroradiometer controlled by the IRIS [12] software. Chromaticities of stimuli were defined in an opponent cone-contrast color space [13,14] (Fig. 1) with the two coordinate axes corresponding to L-M and S cone contrast, respectively [10]. In this color space, chromaticities are defined based on cone contrast with respect to the neutral gray at its center. The distance from the center corresponds to chroma, and the azimuth angle corresponds to hue. The orthogonal third axis corresponds to luminance contrast. Cone contrasts were defined with respect to a neutral gray [64.62 cd/m^2, CIE $(x, y) = (0.317, 0.327)$]. S-cone contrasts were scaled by a factor of 2.6, yielding approximately equally salient stimuli for all hues [15]. Individual perceptual isoluminance with respect to the reference gray was determined for 16 stimuli of different hues using heterochromatic flicker photometry [16]. From these data, an isoluminant plane in cone-opponent color space was calculated for each subject [15]. Stimuli were uniform isoluminant square patches presented in uniform surrounds. Three stimulus sizes, 0.5° × 0.5°, 2° × 2°, or 8° × 8°, were used. Test stimulus chromaticities were regularly spaced along the azimuth of color space at 22.5°, 67.5°, 112.5°, 157.5°, 202.5°, 247.5°, 292.5°, 337.5°, and were moderately saturated with a cone contrast $c = 0.2$ with respect to the gray surround. For the surround, the neutral reference gray and eight chromaticities with cone contrast $c = 0.16$ and azimuth angles rotated by 22.5° from the values for the test stimuli were used (Fig. 1). The display was divided to form two surrounds, each subtending 11° by 17° of visual angle. One surround was neutral gray, and the other was isoluminant with a chromaticity corresponding to one of the surround chromaticities (see Fig. 1). Stimuli were presented in the center of each surround.

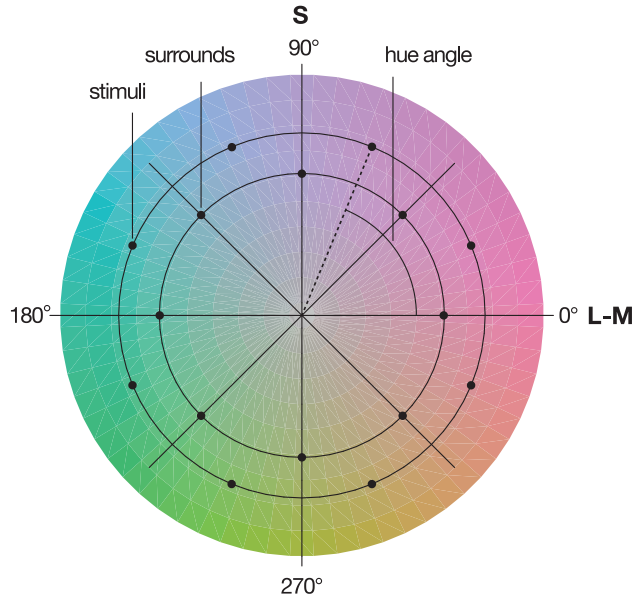


Fig. 1. Isoluminant plane of cone-opponent color space. Horizontal and vertical axes correspond to L-M and S cone contrast, respectively. Stimulus hue is represented as polar angle in this plane. Distance from center corresponds to saturation. Black dots indicate the chromaticities of surrounds and test stimuli.

C. Procedures

During an experimental session, the subject sat in a dimly lit room and viewed the display binocularly from a distance of 1.14 m. Subjects were instructed to maintain their gaze on the screen but were otherwise not restricted in viewing. Experimental sessions were organized in blocks of 64 trials. In each trial, the subject performed asymmetric matching by adjusting the hue of the stimulus in the neutral surround to match it to the test stimulus in the chromatic surround. Whether the test stimulus was presented in the left half of the display and the match stimulus on the right, or vice versa, was selected pseudo-randomly from trial to trial. Hue adjustment was done by moving a computer mouse, which changed the chromaticity along the azimuth of cone-opponent color space. The initial chromaticity of the match stimulus was chosen randomly from an interval of 90° in color space centered around the test hue. After a subject had indicated that a satisfactory match was achieved by pressing the space bar on the keyboard, a 20 Hz flickering checkerboard pattern of chromatic squares with the same size as the stimuli that filled the whole screen was presented for 2 s to prevent afterimages of the stimuli. Chromaticity of each checkerboard square was selected randomly out of the eight test stimulus chromaticities. One block of 64 trials consisted of all combinations of the eight chromatic backgrounds and four of the test stimuli. Stimulus size was constant within each block but varied between blocks. Each observer performed a minimum of 18 experimental sessions, resulting in a total of 1152 trials (8 surrounds × 8 test stimuli × 3 sizes × 6 repetitions). In addition, subjects performed separate sessions with a control condition where both test and match stimuli were displayed in neutral gray surrounds.

D. Analysis

Hue changes were determined as the difference in hue angle between the stimulus displayed in the chromatic surround and the matching chromaticity set by the subject in the neutral surround. For each stimulus size separately, these differences were plotted as a function of angular distance between the chromaticity of the stimulus and the chromaticity of the respective surround.

To quantify the influence of stimulus size on induction strength, the *spread* of induced changes across stimulus sizes was determined. The spread was calculated for each combination of stimulus hue and surround as the difference between the largest and the smallest induced hue shift that were elicited due to different stimulus sizes. The location of maximum spread was then determined for each surround, and the absolute induced hue changes at that location were plotted as a function of stimulus size. How maximum spread depends on the specific surround was further illustrated by plotting the maximal induced spread as a function of surround in a polar plot. An ellipse was fitted to determine the axes of largest and smallest size dependence. The aspect ratio between the major axis a and the minor axis b was used as an intuitive measure of how dissimilar the resulting ellipse was from a circle ($r = \frac{a}{b}$, where $r = 1$ corresponds to a circle).

Testing whether two hue induction curves were significantly different from each other was done via a variation of the chi-squared (χ^2) test. For two induced hue shift curves $f(h)$ and

$g(h)$, measured at hue angles h with standard errors $s_f(h)$ and $s_g(h)$, respectively, the approximate $\tilde{\chi}^2$ value was calculated as

$$\tilde{\chi}^2 = \sum_h \frac{(f(h) - g(h))^2}{s_f^2(h) + s_g^2(h)}.$$

A significant difference between the two induction curves $f(h)$ and $g(h)$ was attested when the probability $p = P(\chi^2 > \tilde{\chi}^2)$ was lower than the chosen significance level α.

Software packages used for the data analysis included Python (2.7.10), NumPy [17] (1.10.1), and Pandas [18] (0.14.1). Plots were generated with Matplotlib [19] (1.4.3). All data and scripts necessary to generate the figures of this paper are publicly available via the German Neuroinformatics Node [20].

3. RESULTS

The induced hue shifts, when plotted as a function of the hue difference between stimulus and surround, showed the same characteristic dependence that was found in previous experiments [3]. For all chromatic surrounds and stimulus sizes, chromatic induction shifted the perceived hue away from the surround hue, thus enlarging their perceptual difference. The condition with 2° test stimulus size in the present study replicated the experiment done by Klauke and Wachtler [3]. Since one subject (TW) took part in both studies, we compared the induced hue shifts measured in both studies. For all surrounds, induced shift changes matched those in the previous study, and none of the curves were significantly different (χ^2 significance test, significance level $\alpha = 0.01$, the p-values are 0°: 0.128, 45°: 0.577, 90°: 0.629, 135°: 0.05, 180°: 0.11, 225°: 0.067, 270°: 0.029, 315°: 0.251). Thus, although most of the hardware including computer and graphics card as well as the software for calibration and stimulus generation was different, the results could be reproduced with high reliability.

It should be noted that the restriction of stimulus adjustment to variation along the azimuth may systematically affect the resulting hue shift values. Assuming that the actual induced color shift is along the line in color space connecting stimulus and surround, the amount of shift indicated by the subject would be expected to be smaller due to the projection of the actual shift onto the circle of constant chroma defined by the stimulus chromaticities. This reduction is a monotonously decreasing function of the difference between stimulus and surround and thus, while affecting the exact shape of the resulting curves of induced hue shifts, cannot be an explanation for their overall shape. Likewise, previous analysis has shown that the curves of induced hue shift also differ systematically from what would be expected if the effect of induction would correspond to a simple shift or compression in color space [3].

While for a given chromatic surround the induced hue changes showed the same qualitative dependence on the hue difference between stimulus and surround, the magnitude of the induced hue changes strongly depended on stimulus size. Figure 2 shows induced hue changes for one surround (0°) for all subjects and stimulus sizes. While the curves for different subjects are similar and cluster for each condition, there are systematic differences for different stimulus sizes. Smaller stimulus sizes led to larger hue changes. For example, the peak

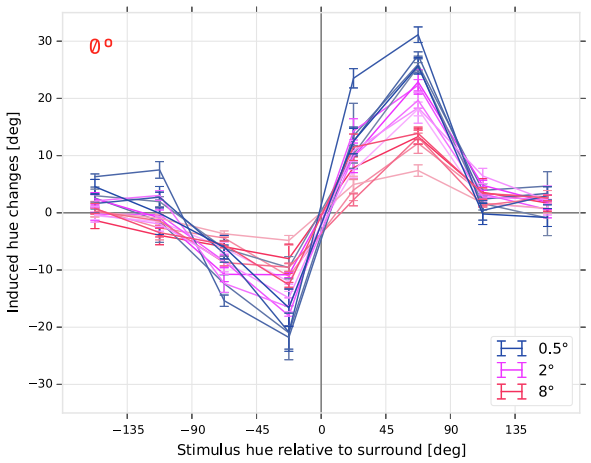


Fig. 2. Hue changes induced by a chromatic surround. Data of five subjects for one of the surrounds (0°) and all three stimulus patch sizes (0.5°, 2°, 8°). Data points are averages of at least six presentations for every stimulus size each. Induced hue changes are plotted as a function of the angular distance between the hue (azimuth angle) of the test stimulus and the hue of the chromatic surround. Color encodes stimulus patch size (different hue) and subject (different saturation). Bars denote standard error.

values in the curve shown in Fig. 2 varied between 31.14° ± 1.4° for the smallest stimulus size of 0.5° and 7.39° ± 1.0° for the largest stimulus size of 8°.

For other surround hues, the results were similar. Figure 3 depicts hue shifts for eight different surrounds and three patch sizes, averaged over all subjects. Induction curves for all sizes and surrounds were similar in shape and all showed the same pattern of first increasing and then decreasing induced hue shifts. The maximal induced hue changes as well as the differences in hue shifts due to different sizes varied with surround hue. Overall, curves for different sizes for a given surround were more similar in shape than curves for different surrounds. Curves for different stimulus sizes of the same surround were significantly different from each other (pairwise χ^2 test, $\alpha = 0.01$), with the exception of the curves for 2° and 8° for the 270° surround ($p = 0.023$). Stimuli that induced the maximal hue change for a given surround also showed the largest size dependent changes, with the exception of the 315° surround, where the maximal induced hue shift and maximal size difference were neighboring points. The control condition, in which both surrounds were neutral gray, confirmed that subjects were able to accurately perform the match and that stimulus patch size had no significant influence on matching performance (pairwise χ^2 test, $\alpha = 0.01$, all $p \geq 0.33$).

We quantified the size dependence of hue shifts for each combination of stimulus and surround by calculating the spread of the values due to varying stimulus size as the difference between the largest and smallest absolute hue shifts and determining the maximum of these spreads for each curve. Each surround had a distinct location of maximal spread (indicated by an asterisk in Fig. 3), typically at or near the location of the largest induced shift. There was also a correlation ($r = 0.758$, $p = 0.029$) between maximum induced shift and maximum spread. To quantify the dependence of hue shifts on

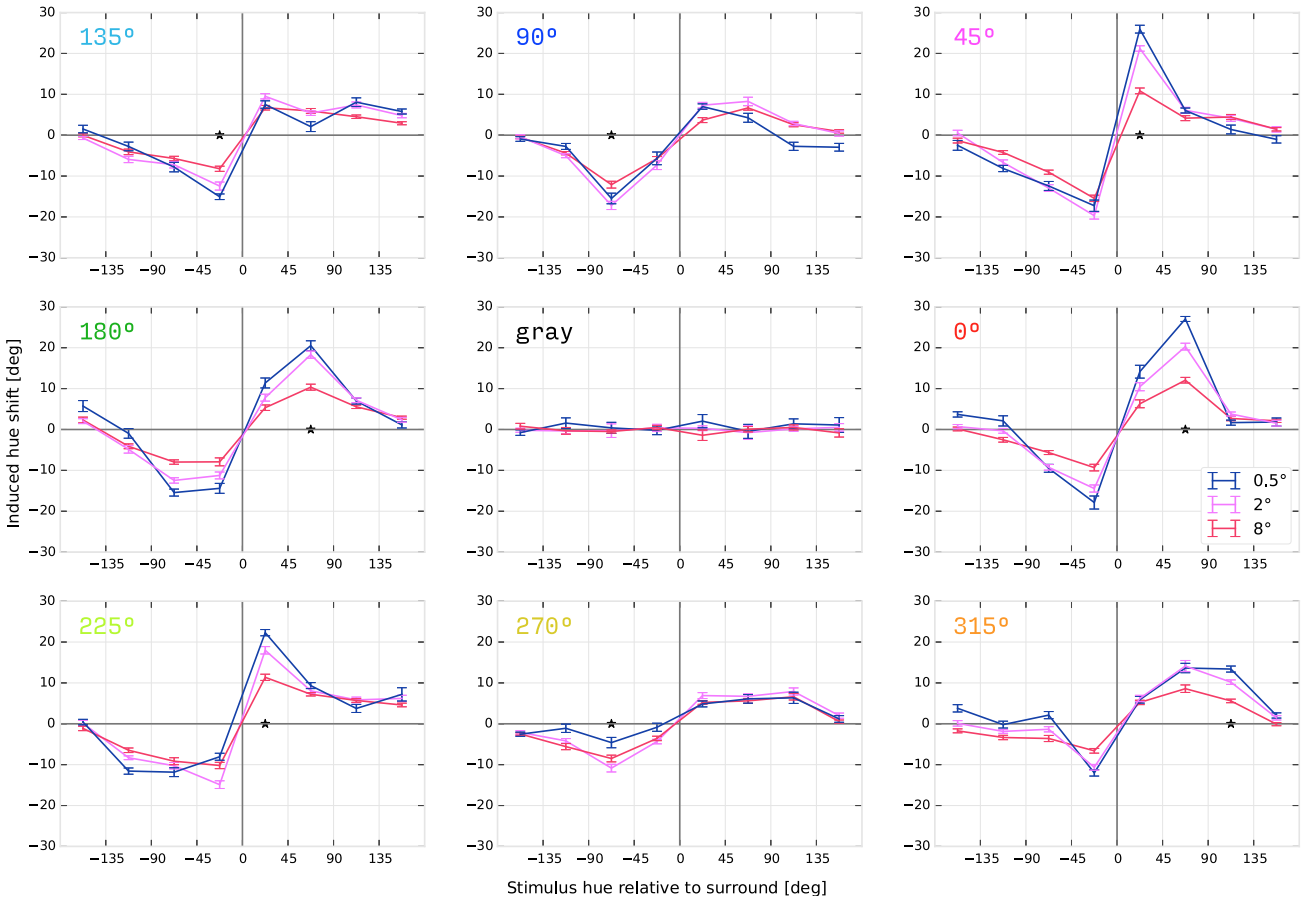


Fig. 3. Induced hue changes for three stimulus sizes and different chromatic surrounds. Each subplot shows the induced hue shifts for all three test stimulus sizes as a function of the angular distance between the chromaticity of the test stimulus and the chromaticity of the surround. Individual data points are averages across five subjects, bars indicate standard error, and the color of the curve encodes stimulus patch size. Different subplots show results for different surrounds and are arranged to reflect their position in color space. Stimulus hues that elicited the strongest size dependencies are marked with an asterisk. The center panel shows data for the control condition, where the subject matched patches of different sizes on neutral gray surrounds.

size for each surround, we plotted the induced shifts for all three sizes at the location of maximal spread. Figure 4 shows these absolute induced hue shifts as a function of test patch size for all eight surrounds. The overall magnitudes of hue shifts differed between curves, reflecting the dependence of induction strength on surround hue, with values varying from below 10° to above 25°. In general, induction decreased with increasing stimulus size, and the relative decrease per degree of stimulus size was in most cases lower for larger stimulus sizes. Shi *et al.* [11] described a logarithmic dependence on stimulus size for brightness induction in achromatic stimuli. Figure 4 plots the size dependence as a function of stimulus size on a logarithmic scale. On this scale, the curves are steeper for the larger stimuli, indicating a reduction of induction strength with stimulus size that is slightly stronger than logarithmic.

Exceptions to the general pattern of size dependence were the curves for surrounds of 90° and 270°, i.e., the cases of induction by a pure (positive or negative) S cone contrast surround on a stimulus defined by L-M contrast. In these cases, induction was weaker for the smallest stimulus sizes of 0.5° than for 2° stimuli. Similar effects can be seen for certain stimuli in the curves for the adjacent 225° and 315° surrounds, where in the curve parts toward 270° have reduced and even at some points no size effect (Fig. 3).

To illustrate the overall differences in size dependence for different surround hues, we plotted the slopes of the curves from Fig. 4 for the larger stimulus sizes between 2° and 8° in a polar plot representing surround hue [Fig. 5(a)]. The plot shows larger size dependence for surrounds in the first and third quadrants than surrounds in the second and fourth quadrants. Fitting an ellipse to the data yielded an inclination of the major axis of 23.19°, which matches the results of fitting ellipses to the induction strengths for fixed stimulus sizes [3]. Thus, the stronger the induction by a surround, the stronger the variation of induction with stimulus size. This suggests that stimulus size modulates induction strength in a multiplicative way. Therefore, we replotted the slopes after scaling the data to the value for the 2° stimulus size [Fig. 5(b)]. The range of slopes on a logarithmic size scale was between 0.16 and 0.32, indicating that doubling the stimulus size reduced the induced hue shift by 16%–32%. An ellipse fit yielded an inclination that was closer to horizontal (11.59° with respect to the L-M axis) and an aspect ratio ($r = 1.43$) that was strongly reduced compared to the ellipse for the unscaled data [$r = 3.75$, Fig. 5],

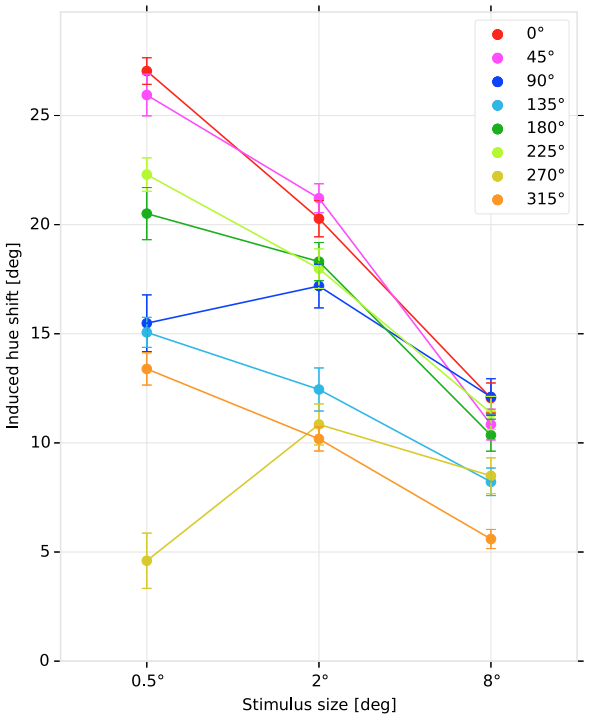


Fig. 4. Stimulus size dependence of induced hue shifts. Induced hue shifts for each surround as a function of stimulus size at the location of maximum spread. Induction is weaker for larger stimuli. Spread was calculated as the distance of the smallest to the largest hue shift at every stimulus hue for each surround separately. Color encodes surround chromaticity. Error bars denote standard error.

indicating that the relative change in induction with stimulus size was similar for the different surrounds. The two surrounds with the lowest slopes were those corresponding to the vertical axis with positive and negative S-cone contrast, respectively.

The separability of surround chromaticity dependence of induction and its modulation by stimulus size indicates that the mechanism that is responsible for the modulation of induction strength by surrounds is independent of the one creating size-related changes. While the bias to an oblique direction in color space [Fig. 5(a)] suggests an adaptation to the properties of the

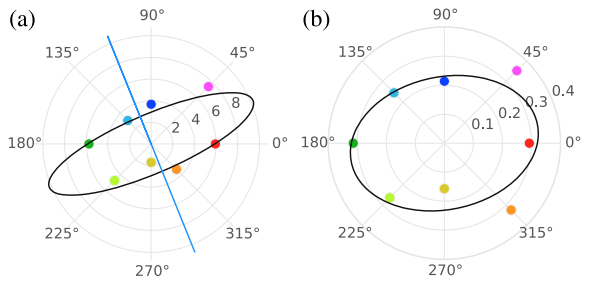


Fig. 5. Size dependence of induction from different surrounds. Polar plots show the (a) absolute and (b) relative slopes of the 2° to 8° hue change curves at the location of maximum spread as a function of the surround hue angle. Best fitting ellipses indicate the directions of surround chromaticities with the largest and smallest size dependence. The relative slopes in (b) were obtained by normalizing the 2°–8° hue change curves to hue shift at 2°. The solid blue line in (a) indicates the perceptual blue–yellow axis.

natural environment and a strong relation to perception [3], the color specificity of the size dependence seems to be coupled to the cardinal axes of opponency and thus to more low-level mechanisms.

4. DISCUSSION

The data obtained in our experiments agree with the results of previous studies investigating chromatic induction [1,3]. Here we measured this induction for stimuli with different sizes and found that there was a systematic modulation of the induction effect with stimulus size. In general, induction strength decreased with stimulus size, so that hue shifts induced by larger stimuli tended to be smaller. This was true with most of the surrounds and indeed with all surrounds for stimulus sizes of 2° and larger.

In their study on the spatial extent of chromatic induction, Brenner and Cornelissen [1] used different stimulus sizes. For stimulus conditions with very small or no gap between stimulus and surround, they reported systematic changes in induction for different stimulus sizes. Our data, obtained with zero gap between stimulus and surround, confirm these observations and moreover systematically quantify this stimulus size dependence of induction.

The effect of stimulus width on simultaneous contrast was recently investigated in a study by Shi *et al.* [11]. Brightness perception was measured for achromatic stimuli of different sizes ranging between 0.25° to 8° of visual angle. Analogous to our data for chromatic stimuli, the authors found a decrease in brightness enhancement with larger stimulus size. They reported that the dependence of induction on stimulus size was well described by a logarithmic function. To a first approximation, this is in line with our data; however, in our data there was a slightly steeper than logarithmic decline in induction strength for large stimulus sizes.

Size dependence varied systematically for different surrounds, showing the same dependence on surround hue as induction strength, with a strong bias toward an oblique axis in color space that coincides with the axis of variation of natural daylight and the perceptual blue–yellow axis [Fig. 5(a)]. Normalizing the data by induction strength to isolate the size effect strongly reduced the bias toward oblique directions in color space and instead revealed a pronounced reduction in size dependence for surrounds along the S cardinal axis. Thus, the increase in induction with decreasing stimulus size was much weaker for S-cone surrounds than for L-M cone surrounds. A possible reason for this difference could be the lower sensitivity for small stimuli in S-cone pathways. Sampling of the retinal image by S cones is relatively sparse, in particular in the fovea [16], yielding poor spatial resolution for S-cone stimuli. Generally, receptive field sizes are larger [21] and sensitivity for high spatial frequencies is much lower in neurons encoding S-cone signals than in other color-selective neurons [22]. Likewise, induction from S-cone signals may be weak at high spatial frequencies. A related phenomenon is that the transition from contrast to assimilation occurs for S-cone inducers at lower spatial frequencies than for L-M inducers [23]. The reduction of hue shifts for the smallest stimulus sizes for S-cone surrounds in our data may be a consequence of this transition.

Induction decreased for larger stimulus sizes, but even for stimulus sizes as large as 8° there was a substantial induction effect. Moreover, even the large stimuli appeared uniform despite the strong hue difference induced by the surround. Usually, filling-in processes are assumed to explain lateral interactions with such spatial ranges [24]. The induction effect is assumed to be initiated at the border, as a result of contrast enhancement by spatially antagonistic receptive fields, and then propagated from the border to fill-in the stimulus surface uniformly. However, at early visual stages the spatial range of lateral interactions is limited [10,25] and typically does not reach the spatial range of lateral interactions in perception [1,26,27]. Moreover, recent findings indicate that at the neural level there is no filling-in for color stimuli in the primary visual cortex [28]. This speaks in favor of a contribution of higher cortical areas to the observed induction effects [8]. In contrast, the finding of a specificity of the size dependence for the S-cone pathway indicates a contribution of early visual stages where S and L-M pathways are separated. Together, these results support the notion that contextual interactions at multiple stages contribute to the processing of sensory signals to form our percept.

Funding. Bundesministerium für Bildung und Forschung (BMBF) (01GQ1004A); Munich Graduate School of Systemic Neurosciences (GSN).

Acknowledgment. The authors thank Delwen Franzen and the two anonymous reviewers for their valuable comments that helped to improve the paper.

REFERENCES

1. E. Brenner and F. W. Cornelissen, "Spatial interactions in color vision depend on distances between boundaries," Naturwissenschaften **78**, 70–73 (1991).
2. Q. Zaidi, B. Yoshimi, N. Flanigan, and A. Canova, "Lateral interactions within color mechanism in simultaneous induced contrast," Vis. Res. **32**, 1695–1707 (1992).
3. S. Klauke and T. Wachtler, ""Tilt" in color space: hue changes induced by chromatic surrounds," J. Vis. **15**(13), 17, 1–11 (2015).
4. D. Jameson and L. M. Hurvich, "Opponent chromatic induction: experimental evaluation and theoretical account," J. Opt. Soc. Am. A **51**, 46–53 (1961).
5. J. M. Eichengreen, "Unique hue loci: induced shifts with complementary surrounds," Vis. Res. **16**, 199–203 (1976).
6. C. Ware and W. B. Cowan, "Changes in perceived color due to chromatic interactions," Vis. Res. **22**, 1353–1362 (1982).
7. V. Ekroll, F. Faul, and R. Niederée, "The peculiar nature of simultaneous colour contrast in uniform surrounds," Vis. Res. **44**, 1765–1786 (2004).
8. A. C. Hurlbert and K. Wolf, "Color contrast: a contributory mechanism to color constancy," Prog. Brain Res. **144**, 147–160 (2004).
9. C. D. Gilbert, J. A. Hirsch, and T. N. Wiesel, "Lateral interactions in visual cortex," Cold Spring Harbor Symp. Quant. Biol. **55**, 663–677 (1990).
10. T. Wachtler, T. J. Sejnowski, and T. D. Albright, "Representation of color stimuli in awake macaque primary visual cortex," Neuron **37**, 681–691 (2003).
11. V. Shi, J. Cui, X. G. Troncoso, S. L. Macknik, and S. Martinez-Conde, "Effect of stimulus width on simultaneous contrast," PeerJ **1**, e146 (2013).
12. C. J. Kellner, "IRIS: toolbox for color vision science (version 0.1.0) software" (2015), https://github.com/wachtlerlab/iris.
13. D. I. MacLeod and R. M. Boynton, "Chromaticity diagram showing cone excitation by stimuli of equal luminance," J. Opt. Soc. Am. A **69**, 1183–1186 (1979).
14. A. M. Derrington, J. Krauskopf, and P. Lennie, "Chromatic mechanisms in lateral geniculate nucleus of macaque," J. Physiol. **357**, 241–265 (1984).
15. H. J. Teufel and C. Wehrhahn, "Evidence for the contribution of S cones to the detection of flicker brightness and red–green," J. Opt. Soc. Am. A **17**, 994–1006 (2000).
16. P. K. Kaiser and R. M. Boynton, *Human Color Vision*, 2nd ed. (Optical Society of America, 1996).
17. S. van der Walt, S. Colbert, and G. Varoquaux, "The numPy array: a structure for efficient numerical computation," Comput. Sci. Eng. **13**, 22–30 (2011).
18. W. McKinney, "Data structures for statistical computing in python," in *Proceedings of the 9th Python in Science Conference*, S. van der Walt and J. Millman, eds. (2010), pp. 51–56.
19. J. Hunter, "Matplotlib: a 2d graphics environment," Comput. Sci. Eng. **9**, 90–95 (2007).
20. C. J. Kellner and T. Wachtler, "Dataset: stimulus size dependence of induced hue changes" (2015), https://dx.doi.org/10.12751/g-node.t6vbz9.
21. C. Tailby, B. A. Szmajda, P. Buzás, B. B. Lee, and P. R. Martin, "Transmission of blue (S) cone signals through the primate lateral geniculate nucleus," J. Physiol. **586**, 5947–5967 (2008).
22. A. J. White, S. G. Solomon, and P. R. Martin, "Spatial properties of koniocellular cells in the lateral geniculate nucleus of the marmoset Callithrix jacchus," J. Physiol. **533**, 519–535 (2001).
23. D. Cao and S. K. Shevell, "Chromatic assimilation: spread light or neural mechanism?" Vis. Res. **45**, 1031–1045 (2005).
24. M. A. Paradiso and K. Nakayama, "Brightness perception and filling-in," Vis. Res. **31**, 1221–1236 (1991).
25. D. M. Alexander and J. J. Wright, "The maximum range and timing of excitatory contextual modulation in monkey primary visual cortex," Vis. Neurosci. **23**, 721–728 (2006).
26. J. Walraven, "Spatial characteristics of chromatic induction; the segregation of lateral effects from straylight artefacts," Vis. Res. **13**, 1739–1753 (1973).
27. T. Wachtler, T. D. Albright, and T. J. Sejnowski, "Nonlocal interactions in color perception: nonlinear processing of chromatic signals from remote inducers," Vis. Res. **41**, 1535–1546 (2001).
28. S. Zweig, G. Zurawel, R. Shapley, and H. Slovin, "Representation of color surfaces in V1: edge enhancement and unfilled holes," J. Neurosci. **35**, 12103–12115 (2015).

5

CONCLUSIONS AND OUTLOOK

How color is encoded and processed in the human retina is now well studied, and even though many open questions remain, the basic coding and information processing is well understood (B. B. Lee, 2011). Much less so for higher stages. The coding in the lateral geniculate nucleus (LGN) is widely believed to be the same as that in the output of the retina. The LGN is therefore often called a relay station, even though only a minority of its connections emanate from the retina while many more come from the cortex and the brainstem (Erişir, Van Horn, & Sherman, 1997; Van Horn, Erişir, & Sherman, 2000; Guillery & Sherman, 2002). As signals reach the cortex, a transformation of the code takes places that is accompanied by a vast increase of available neurons. Through this transformation, color becomes represented by a large population of neurons with broad tuning curves and distributed color preferences. This re-encoding might reflect a change in coding strategy from one that focuses on efficiently transmitting retinal signals through a limited capacity medium (the optic nerve), to one that facilitates further cortical processing. The guiding principles and mechanisms underlying the transformation of the code are poorly understood. In this work we tried to further our understanding of the population code for color in the visual cortex.

In the first study presented in Chapter 2 we used the theoretical framework of *efficient coding*, the idea that signal processing of sensory systems should form an efficient neural code of the world, to study how such a population code for color could arise. For this, we used typical signals from the output stage of the retina. We filtered L-, M- and S-cone signals derived from natural scenes, analogous to the processing of the retina, and performed Independent Component Analysis (ICA) on them to learn an efficient code for such signals. The resulting code was highly sparse and disperse, and color preferences were broadly distributed in colorspace. This confirms the hypothesis that although the opponent code used by subcortical pathways is an efficient representation for unfiltered cone signals (T.-W. Lee et al., 2002), a broadly distributed, sparse code is a more efficient representation for the post-retinal signals. A possible explanation for this is that the center-surround structure of the retina per-

forms a spatial decorrelation on the cone signals (Doi, Inui, Lee, Wachtler, & Sejnowski, 2003; Wachtler, Doi, Lee, & Sejnowski, 2007), increasing the variance in the chromatic channels. Additionally, splitting signals into separate ON and OFF pathways projects the signal into a higher dimensional space, which should enhance feature separation (Schölkopf & Smola, 2002).

The composition of the surround in the pre-filtering step consisted of a mixture of L-, M- and S-cone signals, each contributing equally. Since the exact composition of the surround is still a question of serious debate (Lennie et al., 1991; Reid & Shapley, 1992; B. B. Lee et al., 1998; Martin et al., 2001; Reid & Shapley, 2002; Buzás et al., 2006; G. D. Field et al., 2010; Crook et al., 2011; B. B. Lee et al., 2012), we also tried different surround variants. The mixed surround was the most efficient, both in terms of population kurtosis and lifetime kurtosis[1], but the difference was not big enough to draw any conclusions. Therefore, from a theoretical perspective, this remains an open question.

The considerations regarding the surround compositions naturally lead to another interesting question: What if one of the cone classes is completely removed as a source of information? Most mammals only possess two cone types, one sensitive to the short wavelength part of the spectrum and another sensitive to longer wavelengths. Moreover, in color-blind humans one of the three cone types is missing, reducing the three dimensional input space of humans with normal trichromatic vision to a two dimensional (dichromatic) subset. In dichromats with protanopia the L-cone is missing and in dichromats with deuteranopia the M-cone is absent, resulting in what is normally called "red-green" color blindness. Tritanopia, where S-cones are absent, leads to "blue-yellow" color blindness. When dichromatic vision is simulated for trichromats it appears as if some colors are missing completely (Viénot, Brettel, Ott, Ben M'Barek, & Mollon, 1995; Brettel, Viénot, & Mollon, 1997). Surprisingly though, when asked to report their percepts, color-blind people use all four basic color terms, i.e. "red", "green", "blue", and "yellow" (Boynton & Scheibner, 1967; Scheibner & Boynton, 1968; Wachtler, Dohrmann, & Hertel, 2004). This would suggest that the dimensionality of their perceptual space is higher than two, as implied by the number of available cone types. A possible

1 Briefly, in a code with high *population kurtosis*, or *population sparseness*, only a small number of available neurons is active for a specific stimulus, and also a different subset is active for a different stimulus. This contrasts to a *compact code*, where only a small number of neurons is active for all stimuli. *Lifetime kurtosis* on the other hand, means that a neuron is inactive for most of the time, but if it becomes active its response will be large (Willmore & Tolhurst, 2001).

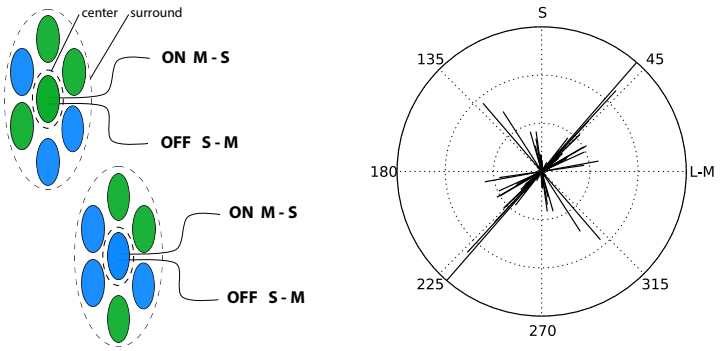


Figure 5.1: Efficient coding for dichromatic natural signals.
LEFT| Schematic of a protanope-like center-surround structure with individual ON, OFF pathways for S-center and M-center cell types.
RIGHT| Estimates of color preferences for basis functions learned from cone protanope-like retinal cone activations. Signals were obtained from natural images that were filtered with a center-surround like kernel and then split into ON and OFF signals.

solution to this ostensible contradiction might lie in the non-linearities that arise through separate ON and OFF channels. Figure 5.1 LEFT shows a schematic representation of a retinal ganglion cell for a protanope-like (missing L-cone) situation with a center-surround structure and the corresponding ON and OFF channels. In a preliminary study, we performed ICA on dichromatic signals that were filtered as in the study of Chapter 2. The analysis of the resulting four dimensional signals is not obvious and the methods used in the previous studies are not completely accurate. Nevertheless, a tentative analysis showed that all basis functions obtained with ICA were color selective and most of them were color-opponent. Color preferences of basis functions exist in a four-dimensional[2] space with no obvious transformation back into the original three-dimensional colorspace. Projections of the four-dimensional cone activation space into a three-dimensional sub-space show that the basis functions seem to span the entire four-dimensional space. This is furthermore illustrated by the wide distribution of basis functions when plotted in the cone-opponent colorspace, as shown in Figure 5.1 RIGHT. These results, particularly the distribution of basis functions, were performed with preliminary analysis methods and should be treated with caution. However, they provide

2 In the case of a protanope-like situation, the four dimensions are S+, S-, M+ and M-, where "+" denotes the ON channel and "-" the OFF counterpart.

a promising outlook that there is enough structure in the statistics of the cone signals of dichromats for them to be able to adequately form the same four basic color categories. More in depth analysis needs to be carried out to confirm the preliminary results.

In the visual system, the *efficient coding* hypothesis has been applied to the retina (e.g. Doi et al., 2012) and explains features of primary visual cortex (e.g. Olshausen and Field, 1996; Bell and Sejnowski, 1997). But what about higher stages? A few studies have extended the approach to V2 (H. Lee, Ekanadham, & Ng, 2008; Shan & Cottrell, 2013; Hosoya & Hyvärinen, 2015) and reproduced properties of neurons found in V2, such as angle selectivity (Ito & Komatsu, 2004). All models are hierarchical in that they first learn V1 receptive fields (orientation selective cells) and then use that as a source to train another sparse model. Hosoya and Hyvärinen (2015), for example, created a sparse code model of V2, where the inputs to the V2 model consisted of the outputs of a primary visual cortex (V1) sparse coding model. The V1 model consisted of two types of units that mimicked simple- and complex-like cells. The "simple cells" were represented by a filter bank with Gabor kernels, as found in previous sparse code models (Olshausen & Field, 1996). These formed the input to a second V1 layer to generate phase invariant V1 complex cell-like units. Such cells also emerged from sparse coding of natural scenes (Hyvärinen & Hoyer, 2000; Hyvärinen & Hoyer, 2001; Karklin & Lewicki, 2009). They performed ICA to learn a sparse code on the outputs of the complex-cell like units, and found that their model units exhibited properties similar to some features of V2 cells that were experimentally reported in macaque monkeys. Other properties could not yet be reproduced, such as the combination of orientations (Anzai, Peng, & Van Essen, 2007) or texture boundary detection (Schmid, Purpura, & Victor, 2014). What about even higher stages in the visual system, where some cells respond to complex object features (Tanaka, 1996, 2003), faces and body parts (Pinsk, DeSimone, Moore, Gross, & Kastner, 2005; Tsao, Freiwald, Tootell, & Livingstone, 2006), single persons (Quiroga et al., 2005) and even scenes (Kornblith, Cheng, Ohayon, & Tsao, 2013)? It is very likely that other mechanisms not explainable with redundancy reduction and efficient coding alone are at work in the visual system; mechanisms that are linked to the goal of the visual system.

One of the main criticisms[3] of the *efficient coding hypothesis* is that it is only concerned with encoding and the constraints of neurons, and is

3 See the review from Simoncelli (2003) for a more comprehensive list of criticisms together with a large list of experiments and their relation and contributions to the *efficient coding hypothesis*.

oblivious to the task that the organism must perform and therefore also ignores the task of its sensory system (Simoncelli & Olshausen, 2001; Simoncelli, 2003). This might be closely related to the more complex features of higher stage neurons mentioned in the previous paragraph. Another, perhaps even more fundamental criticism, is that the input and output are not properly defined within the framework (Simoncelli, 2003). As the example of a single neuron from Chapter 1.1.2 illustrates, efficient coding depends heavily on the input, coding scheme, and constraints of the output. Therefore, the whole theory hinges on accurate estimations of natural statistics and on the way in which neurons code information. Especially the latter — whether a neuron encodes its information via its spike rate, spike timing, or some other measure — is often not certain. Moreover, in depth quantitative analyses of redundancy related performance measures ("multi-information", "average log-loss", "rate–distortion") have cast doubt on the finding that a linear filter bank with V1-like simple cells is significantly better in removing said redundancies compared to decorrelation methods such as Principal Component Analysis (PCA). This challenges the conclusion that orientation selective filters are good at capturing higher order correlations (Bethge, 2006; Eichhorn et al., 2009).

In summary, the *efficient coding hypothesis* has been quite influential, has stipulated a lot of research and was applied successfully to many aspects of vision and also to other modalities, such as the auditory system (Lewicki, 2002; Chechik et al., 2001). Extensions of the theory, for example with "active perception" (Lonini et al., 2013) or related frameworks, that make similar assumptions and take the statistics of the environment into account, like Bayesian theories (e.g. Doya, Ishii, Pouget, and Rao, 2007; Pouget, Beck, Ma, and Latham, 2013) seem promising. Further research will show how far the hypothesis can be used to explain mechanisms of sensory systems and the brain, or in turn identify its limitations.

In Chapter 3 we describe a population code model for color that can quantitatively predict human hue induction effects. Such induction effects occur when a chromatic stimulus is presented on a colored surround. This leads to a systematic change in the perception of the stimulus hue, i.e. the surround induces a perceptual change of the stimulus hue. The strength of the hue shift depends on two factors: the distance in colorspace between the stimulus and surround hue, and on the specific surround hue. The model consists of color selective units with bell-shaped, broad and overlapping tuning curves, whose color preferences are uniformly distributed around the hue circle. The contextual effect was implemented as inhibitory modulation following a Gaussian profile centered at the

surround hue, thereby reducing the model units' activity. Therefore, the closer a unit is to the background, the higher its reduction and the lower its activity. There were two free parameters: the strength of the modulation, corresponding to the maximum of the Gaussian, and the modulation width, corresponding to its standard deviation. The modulation led to a change in the pattern of activity in the population code that, in turn, resulted in a change of the encoded hue. By varying these two parameters, we could adequately reproduce all human hue induction curves for different backgrounds.

At the physiological level, contextual modulation in sensory systems is closely related to the concept of non-classical receptive fields (nCRF). If a classical receptive field (CRF) is the subregion of the input stimulus space that leads to a direct response, then the nCRF is the subregion of the stimulus space that can itself not evoke a response but can modulate a response (Allman, Miezin, & McGuinness, 1985). Wachtler et al. (2003) found that presenting a chromatic stimulus in the nCRF of color-selective neurons in primary visual cortex of macaque monkeys changed their response in a systematic way that is compatible with human hue induction effects. Thus it seems very likely that neurons in V1 are involved in color appearance and perceptual effects such as hue induction. These findings partially motivated our model in Chapter 3.

There is an interesting link between efficient coding and contextual modulation. The influence of the non-classical receptive field on coding efficiency and sparseness in primary visual cortex has been studied in awake and behaving monkeys (Vinje & Gallant, 2000; Vinje & Gallant, 2002). They created movies from grayscale natural images by mimicking typical saccadic eye movements around the image. Along the path of the movements they sampled image patches of different sizes up to four times the size of the classical receptive field. The result was that the generated movie that was presented in the CRF was always the same, but the extent of the stimulation in the nCRF varied. The presence of stimuli in the nCRF led to either the enhancement or suppression of responses — the latter being the predominant effect — which, in turn, decreased the population activity. For single neurons it led to a higher "information spike" and "information rate", and increased the "efficiency of information transmission" (Vinje & Gallant, 2002).

Felsen, Touryan, and Dan (2005) investigated the effect of contextual modulation on coding efficiency in a population code model. They focused on orientation in natural scenes and examined the spatial and temporal statistics in natural images and movies. They modeled orientation detec-

tors (Adelson & Bergen, 1985) and found that the changes of predominant orientations in time (for movies) or location (for images) were likely to be zero or small. Recording from orientation selective cells in the primary visual cortex of cats revealed that showing gratings with the same orientation in subsequent frames modulates the tuning of most neurons by shifting their tuning away from the preferred orientation. They then modeled a population of orientation selective units with identical, overlapping and Gaussian-like tuning curves whose preferred orientation was uniformly distributed, very similar to our population code in Chapter 3. While in our model the spatial surround had an inhibitory effect on the response, the model of Felsen et al. (2005) shifted the tuning curves of the units. The shift depended on the angular difference in orientation and followed a Gaussian profile, with its mean (i.e. the maximal shift) centered at the surround orientation. The fact that the model tuning curves overlapped implies a degree of inherent redundancy in the population code, since nearby units will respond in a similar way to the same orientation. The repulsive shift changed the distribution of tuning curves and their density, with two consequences: on the one hand, it decreased the overlap and therefore redundancy for units whose preferred tuning was close to the surround orientation; on the other hand, it increased the overlap and thus redundancy in regions whose orientation was further away from the surround. Since in natural images and movies nearby pixels or frames are more likely to be of similar orientation, the contextual modulation resulted in an overall redundancy reduction of up to 33%.

The architecture of the population code model in Chapter 3 is quite simplistic. This was intentional since we didn't want to introduce too many free parameters. It is very likely that some of the features found in the hue induction curves, such as the offset in the zero crossing and the shape of the peaks, could be captured by the model if tuning curves were more realistic. In the psychophysics experiment, the surround was of slightly different saturation in order to be able to tell center and surround apart when they had the same hue. This might explain the shift of the zero crossing. A model with units that are also tuned to saturation would be necessary to test this hypothesis. In general, the tuning curves for color found in physiological experiments are quite diverse. Wachtler et al. (2003), for example, measured tuning curves with asymmetric shapes, with tuning widths ranging from around 30° up to 80° (Figure 1.6a). Their preferred tuning was also non-uniformly distributed in colorspace (Figure 1.6b).[4] This all indicates that the actual population code is anisotropic, and that adopting a more realistic model, for example by using existing

4 Although the data from Wachtler et al. (2003) is taken here as an example, see also Lennie et al. (1990), De Valois et al. (2000), Hanazawa et al. (2000).

data for tuning width and peak distributions, could explain more of the features observed in the psychophysics experiments. Moreover, a single, non-surround dependent modulation profile, when applied to an anisotropic population code, could lead to the observed per-surround differences.

The hue encoded in the population code is inseparably linked to the decoding scheme. This is a problem since the actual decoding scheme used by the visual cortex is not known. The decoding scheme we used tried to accommodate for that by making as few assumptions as possible. The only assumption of our decoder is that downstream neurons are "ignorant" of the surround modulation, and should therefore decode similar activity patterns as similar hues even if they resulted from a modulated activation pattern. There are several other established decoding schemes (Pouget & Latham, 2007). The standard decoding scheme is the population vector (Georgopoulos et al., 1986), where the encoded hue is the vector sum of the units' preferred direction[5], weighted with their activity. Such a coding scheme yields interesting results when the population code is anisotropic, because the vector sum does not take density into account. A region with higher density will contribute more to the resulting sum which, in turn, will lead to a bias. This has recently been used to implement Bayesian inference (Rich, Cazettes, Wang, Peña, & Fischer, 2015; Girshick, Landy, & Simoncelli, 2011; Fischer & Peña, 2011): the shape of the tuning curve encodes the likelihood and the density of the tuning curves encode the prior. Girshick et al. (2011) used this to model the oblique effect (Appelle, 1972): human observers are better at discriminating orientations for cardinal axes (horizontal or vertical) than oblique axes. In their population code model of orientation selective units, areas around the cardinal axes had a higher density of tuning curve slopes. If indeed a non-uniform population code is used to encode color in the human brain, biases for preferred colors might also exist, and could be revealed by psychophysics experiments.

An open question in our model is the mechanism underlying surface perception in the presence of surround modulation. Contextual modulation itself is widely believed to be generated by lateral interactions in the visual system that are localized to the border of color edges, followed by a filling-in process (Walraven, 1973; Brenner & Cornelissen, 1991; Paradiso & Nakayama, 1991; Alexander & Wright, 2006) with the involvement of higher stages (Zweig, Zurawel, Shapley, & Slovin, 2015).

5 Direction could be movement direction for motor commands, or orientation for a population code of orientation selective cells, or color in a distributed code for color.

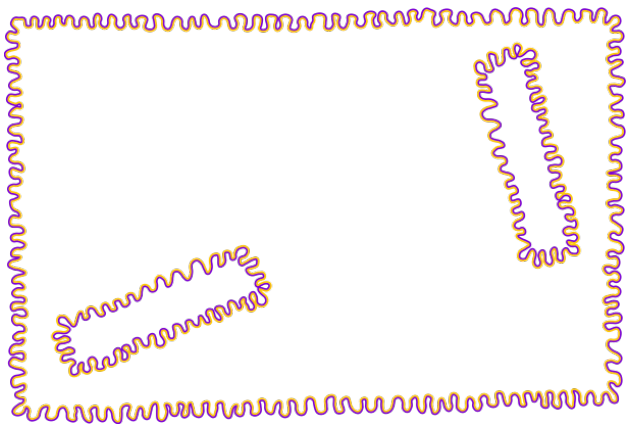
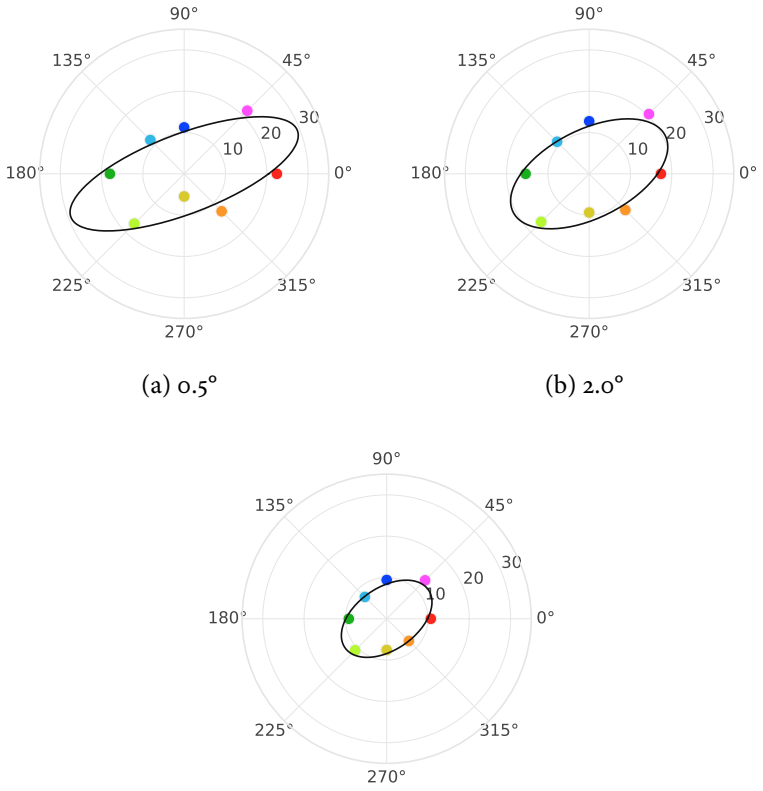


Figure 5.2: Watercolor illusion by Baingio Pinna used under Creative Commons license CC BY-NC-SA.

The watercolor illusion (Pinna, Brelstaff, and Spillmann, 2001, see Figure 5.2), where the existence of a colored border leads to a perceptual filling-in of a white surface with color, suggests that two (or more) distinct steps are involved (Von der Heydt & Pierson, 2006). If so, the influence of stimulus size is a relevant question to be resolved in models of contextual modulation.

In Chapter 4 we address the question of whether stimulus size influences the strength of the context of hue induction. We repeated the hue induction experiments from Klauke and Wachtler (2015), where subjects had to perform asymmetric color matching for colored stimuli in chromatic surrounds, but used three different sizes for the stimulus test patches, ranging from 0.5° to 8°. In general, we found that for almost all backgrounds the induction effect was larger for small stimuli and smaller for large stimuli. Medium size patches induced hue shifts falling between the two. Hue induction therefore does not only depend on the hue difference at the chromatic border, but also depends on the size of the stimulus surface.

The magnitude and general shape of the induced hue shift varied systematically for different surround hues, as in the results from Klauke and Wachtler (2015). For all sizes the induction effects were strongest along an oblique axis in colorspace. An ellipse was fitted to estimate the direction of the oblique axis, for each size separately. The orientation of the ellipses' major axes, with respect to the L-M axis, were 21° for 0.5°, 26.7° for 2° and 34.8° for 8°. Since induction effects were perpendicular to the surround hue, this means that the strongest induction effects are along axes 112°, 116.7° and 124.8°. All these values revolve around the direction of the



(a) 0.5°

(b) 2.0°

(c) 8.0°

Figure 5.3: Maximum induced hue shifts (in degree) as a function of surround hue for different stimulus patch sizes. Data from Kellner and Wachtler, 2016.

cerulean line (112°), which indicates the variation of natural daylights and coincides with the perceptual blue-yellow axis (Mollon, 2006). Along this line, there is a high variation of natural spectra (Webster & Mollon, 1997) and therefore cone activity (Wachtler et al., 2001), clusters of cell preferences in visual cortex (Wachtler et al., 2003) and a lower discrimination performance (Danilova & Mollon, 2009; Witzel & Gegenfurtner, 2013). Another prominent feature of perception, the unique hues, also does not coincide with any of the axes (Valberg, 2001; Webster, Miyahara, Malkoc, & Raker, 2000b, 2000a; Wuerger, Atkinson, & Cropper, 2005). All of this highlights the special role of the daylight axis and indicates that the colorspace is not perceptually uniform. This also leads to the general question of how the cone-opponent colorspace based on retina physiology is linked to perception.

It is possible that non-uniformity is linked to the shape of the ellipse formed by the hue induction maxima, and that the surround specific hue induction effects are a result of the colorspace (Klauke & Wachtler, 2015). If this is indeed the case, a transformation of the colorspace, such as a compression of the major axis, should exist that would make the colorspace more perceptually uniform and thus would increase the similarity of the measured hue induction curves. Initial attempts to find such a transformation were not conclusive, but this might be related to the poor sampling of the space (only eight data points for the eight different surrounds). A psychophysical experiment that measures just noticeable differences along the hue circle in the isoluminant plane could provide the necessary basis to determine the transformation in question. If found, it could provide a clue as to how retinal signals are re-encoded in the cortex and could be used to lay out the tuning curves in a model with an anisotropic population code. As mentioned above, in such a model uniform surround modulation would lead to per-surround hue induction curves of different shapes and magnitudes. Therefore this transformation would constitute a valuable piece of information to solve "one of the central mysteries of colour science" (Mollon & Jordan, 1997), i.e. how retinal mechanisms are related to cortical coding of color and perception.

During the course of the work presented in this thesis, we tried to develop a better understanding of the coding mechanisms of color used by the visual system. Our studies focus on modeling how color is encoded in the early visual cortex, and how contextual modulations of hue perception can be realized within such a model. I hope the research presented here will contribute to our understanding of how color is processed in the brain. Of course new questions arose too, and many mysteries remain.

Adelson, E. H. & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *J. Opt. Soc. Am. A*, *2*(2), 284–299.

Alexander, D. M. & Wright, J. J. (2006). The maximum range and timing of excitatory contextual modulation in monkey primary visual cortex. *Vis. Neurosci. 23*(5), 721–728.

Allman, J., Miezin, F., & McGuinness, E. (1985). Stimulus specific responses from beyond the classical receptive field: neurophysiological mechanisms for local-global comparisons in visual neurons. *Annu. Rev. Neurosci. 8*(1), 407–430.

Angelucci, A. & Sainsbury, K. (2006). Contribution of feedforward thalamic afferents and corticogeniculate feedback to the spatial summation area of macaque V1 and LGN. *J. Comp. Neurol. 498*(3), 330–351.

Anzai, A., Peng, X., & Van Essen, D. C. (2007). Neurons in monkey visual area V2 encode combinations of orientations. *Nat. Neurosci. 10*(10), 1313–1321.

Appelle, S. (1972). Perception and discrimination as a function of stimulus orientation: the "oblique effect"in man and animals. *Psychol. Bull. 78*(4), 266–278.

Atick, J. J. (1992). What does the retina know about natural scenes? *Neural Comput. 4*(2), 196–210.

Atick, J. J. & Redlich, A. N. (1990). Towards a theory of early visual processing. *Neural Comput. 2*(3), 308–320.

Atick, J. J. & Redlich, A. N. (1993). Convergent Algorithm for Sensory Receptive Field Development. *Neural Comput.* 45–60.

Attneave, F. (1954). Some informational aspects of visual perception. *Psychol. Rev. 61*(3), 183–193.

Barlow, H. B. (1961). Possible principles underlying the transformations of sensory messages. In W. A. Rosenblith (Ed.), *Sensory communication* (pp. 217–234). Cambridge, Massachusets: MIT Press.

Bell, A. J. & Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural Comput. 7*(6), 1129–1159.

Bell, A. J. & Sejnowski, T. J. (1997). The "independent components"of natural scenes are edge filters. *Vis. Res. 37*(23), 3327–3338.

Bethge, M. (2006). Factorial coding of natural images: how effective are linear models in removing higher-order dependencies? *J. Opt. Soc. Am. A*, *23*(6), 1253–1268.

Blasdel, G. G. & Lund, J. S. (1983). Termination of afferent axons in macaque striate cortex. *J. Neurosci. 3*(7), 1389–1413.

Box, G. E. P. (1979). Robustness in the Strategy of Scientific Model Building. In R. L. Launer & G. N. Wilkinson (Eds.), *Robustness in statistics* (pp. 40–236). Academic Press.

Boycott, B. B., Dowling, J. E., & Kolb, H. (1969). Organization of the Primate Retina: Light Microscopy. *Phil. Trans. R. Soc. B*, *255*(799), 109–184.

Boynton, R. M. (1996). History and current status of a physiologically based system of photometry and colorimetry. *J. Opt. Soc. Am. A*, *13*(8), 1609–1621.

Boynton, R. M. & Scheibner, H. M. O. (1967). On the Perception of red by "red-blin"observers. *Acta Chromatica*, *1*, 1–19.

Brenner, E. & Cornelissen, F. W. (1991). Spatial interactions in color vision depend on distances between boundaries. *Naturwissenschaften*, *78*(2), 70–73.

Brettel, H., Viénot, F., & Mollon, J. D. (1997). Computerized simulation of color appearance for dichromats. *J. Opt. Soc. Am. A*, *14*(10), 2647–2655.

Buchsbaum, G. & Gottschalk, A. (1983). Trichromacy, opponent colours coding and optimum colour information transmission in the retina. *Proc. R. Soc. Lond., B, Biol. Sci. 220*(1218), 89–113.

Buzás, P., Blessing, E. M., Szmajda, B. A., & Martin, P. R. (2006). Specificity of M and L Cone Inputs to Receptive Fields in the Parvocellular Pathway: Random Wiring with Functional Bias. *J. Neurosci. 26*(43), 11148–11161.

Camus, A. (1970). *Lyrical and Critical Essays*. Vintage Books.

Catterjee, S. & Callaway, E. M. (2003). Parallel colour-opponent pathways to primary visual cortex. *Nature*, *426*(6967), 668–671.

Chechik, G., Globerson, A., Tishby, N., Anderson, M. J., Young, E. D., & Nelken, I. (2001). Group Redundancy Measures Reveal Redundancy Reduction in the Auditory Pathway. In *NIPS*.

Churchland, P. S., Koch, C., & Sejnowski, T. J. (1988). What is Computational Neuroscience? 1–12.

Cottaris, N. P. & De Valois, R. L. (1998). Temporal dynamics of chromatic tuning in macaque primary visual cortex. *Nature*, *395*(6705), 896–900.

Crook, J. D., Davenport, C. M., Peterson, B. B., Packer, O. S., Detwiler, P. B., & Dacey, D. M. (2009). Parallel ON and OFF cone bipolar inputs establish spatially coextensive receptive field structure of blue-yellow ganglion cells in primate retina. *J. Neurosci. 29*(26), 8372–8387.

Crook, J. D., Manookin, M. B., Packer, O. S., & Dacey, D. M. (2011). Horizontal cell feedback without cone type-selective inhibition mediates "red-green"color opponency in midget ganglion cells of the primate retina. *J. Neurosci. 31*(5), 1762–1772.

Curcio, C. A., Allen, K. A., Sloan, K. R., Lerea, C. L., Hurley, J. B., Klock, I. B., & Milam, A. H. (1991). Distribution and morphology of human cone photoreceptors stained with anti-blue opsin. *J. Comp. Neurol. 312*(4), 610–624.

Dacey, D. M. (2000). Parallel pathways for spectral coding in primate retina. *Annu. Rev. Neurosci. 23*, 743–775.

Dacey, D. M., Crook, J. D., & Packer, O. S. (2013). Distinct synaptic mechanisms create parallel S-ON and S-OFF color opponent pathways in the primate retina. *Vis. Neurosci. 31*(02), 139–151.

Dacey, D. M., Liao, H.-W., Peterson, B. B., Robinson, F. R., Smith, V. C., Pokorny, J., … Gamlin, P. D. (2005). Melanopsin expressing ganglion cells in primate retina signal colour and irradiance and project to the LGN. *Nature*, *433*(7027), 749–754.

Danilova, M. V. & Mollon, J. D. (2009). Parafoveal color discrimination: A chromaticity locus of enhanced discrimination. *J. Vis. 10*(1), 4–4.

De Valois, R. L. (1960). Color vision mechanisms in the monkey. *J. Gen. Physiol. 43(6)Suppl*(6), 115–128.

De Valois, R. L., Cottaris, N. P., & Elfar, S. D. (2000). Some transformations of color information from lateral geniculate nucleus to striate cortex. *Proc. Natl. Acad. Sci. U.S.A. 97*(9), 4997–5002.

De Valois, R. L., Smith, C. J., Kitai, S. T., & Karoly, A. J. (1958). Response of single cells in monkey lateral geniculate nucleus to monochromatic light. *Science*, *127*(3292), 238–239.

Derrington, A. M., Krauskopf, J., & Lennie, P. (1984). Chromatic mechanisms in lateral geniculate nucleus of macaque. *J. Physiol. (Lond.) 357*, 241–265.

Doi, E., Gauthier, J. L., Field, G. D., Shlens, J., Sher, A., Greschner, M., … Simoncelli, E. P. (2012). Efficient coding of spatial information in the primate retina. *J. Neurosci. 32*(46), 16256–16264.

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 Comput. 15*(2), 397–417.

Doi, E. & Lewicki, M. S. (2014). A simple model of optimal population coding for sensory systems. *PLoS Comput. Biol. 10*(8), e1003761–e1003761.

Dong, D. W. & Atick, J. J. (1995). Statistics of natural time-varying images. *Network*, *6*(3), 345–358.

Doya, K., Ishii, S., Pouget, A., & Rao, R. P. N. (2007). *Bayesian Brain*. Probabilistic Approaches to Neural Coding. Cambridge, MA: MIT Press.

Eichhorn, J., Sinz, F., & Bethge, M. (2009). Natural Image Coding in V1: How Much Use Is Orientation Selectivity? *PLoS Comput. Biol. 5*(4), e1000336.

Erişir, A., Van Horn, S. C., & Sherman, S. M. (1997). Relative numbers of cortical and brainstem inputs to the lateral geniculate nucleus. *Proc. Natl. Acad. Sci. U.S.A. 94*(4), 1517–1520.

Felsen, G., Touryan, J., & Dan, Y. (2005). Contextual modulation of orientation tuning contributes to efficient processing of natural stimuli. *Network*, *16*(2-3), 139–149.

Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *J. Opt. Soc. Am. A*, *4*(12), 2379–2394.

Field, G. D., Gauthier, J. L., Sher, A., Greschner, M., Machado, T. A., Jepson, L. H., … Chichilnisky, E. J. (2010). Functional connectivity in the retina at the resolution of photoreceptors. *Nature*, *467*(7316), 673–677.

Fischer, B. J. & Peña, J. L. (2011). Owl's behavior and neural representation predicted by Bayesian inference. *Nat. Neurosci. 14*(8), 1061–1066.

Gegenfurtner, K. R., Kiper, D. C., & Fenstemaker, S. B. (1996). Processing of color, form, and motion in macaque area V2. *Vis. Neurosci. 13*(1), 161–172.

Gegenfurtner, K. R., Kiper, D. C., & Levitt, J. B. (1997). Functional properties of neurons in macaque area V3. *J. Neurophysiol. 77*(4), 1906–1923.

Gegenfurtner, K. R. (2003). Cortical mechanisms of colour vision. *Nat. Rev. Neurosci. 4*(7), 563–572.

Georgopoulos, A. P., Schwartz, A. B., & Kettner, R. E. (1986). Neuronal population coding of movement direction. *Science*, *233*(4771), 1416–1419.

Gerstner, W., Kreiter, A. K., Markram, H., & Herz, A. V. (1997). Neural codes: firing rates and beyond. *Proc. Natl. Acad. Sci. U.S.A. 94*(24), 12740–12741.

Giocomo, L. M., Moser, M.-B., & Moser, E. I. (2011). Computational Models of Grid Cells. *Neuron*, *71*(4), 589–603.

Girshick, A. R., Landy, M. S., & Simoncelli, E. P. (2011). Cardinal rules: visual orientation perception reflects knowledge of environmental statistics. *Nat. Neurosci. 14*(7), 926–932.

Gross, C. G. (2002). Genealogy of the "grandmother cell". *Neuroscientist*, *8*(5), 512–518.

Grothe, B. (2003). New roles for synaptic inhibition in sound localization. *Nat. Rev. Neurosci. 4*(7), 540–550.

Guillery, R. W. & Sherman, S. M. (2002). Thalamic Relay Functions and Their Role in Corticocortical Communication. *Neuron*, *33*(2), 163–175.

Hanazawa, A., Komatsu, H., & Murakami, I. (2000). Neural selectivity for hue and saturation of colour in the primary visual cortex of the monkey. *Eur. J. Neurosci. 12*(5), 1753–1763.

Hartline, H. K. (1938). The response of single optic nerve fibers of the vertebrate eye to illumination of the retina. *Am. J. Physiol. 121*(2), 400–421.

Hartline, H. K. (1969). Visual Receptors and Retinal Interaction. *Science*, *164*(3877), 270–278.

Henderickson, A. E., Wilson, J. R., & Ogren, M. P. (1978). The neurological organization of pathways between the dorsal lateral geniculate nucleus and visual cortex in old world and new world primates. *J. Comp. Neurol. 182*(1), 123–136.

Hendry, S. H. & Yoshioka, T. (1994). A neurochemically distinct third channel in the macaque dorsal lateral geniculate nucleus. *Science*, *264*(5158), 575–577.

Horwitz, G. D., Chichilnisky, E. J., & Albright, T. D. (2005). Blue-Yellow Signals Are Enhanced by Spatiotemporal Luminance Contrast in Macaque V1. *J. Neurophysiol. 93*(4), 2263–2278.

Horwitz, G. D., Chichilnisky, E. J., & Albright, T. D. (2007). Cone Inputs to Simple and Complex Cells in V1 of Awake Macaque. *J. Neurophysiol. 97*(4), 3070–3081.

Hosoya, H. & Hyvärinen, A. (2015). A Hierarchical Statistical Model of Natural Images Explains Tuning Properties in V2. *J. Neurosci. 35*(29), 10412–10428.

Hubel, D. H. & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol. (Lond.) 160*, 106–154.

Hyvärinen, A. & Hoyer, P. (2000). Emergence of phase- and shift-invariant features by decomposition of natural images into independent feature subspaces. *Neural Comput. 12*(7), 1705–1720.

Hyvärinen, A. & Hoyer, P. O. (2001). A two-layer sparse coding model learns simple and complex cell receptive fields and topography from natural images. *Vis. Res. 41*(18), 2413–2423.

Ito, M. & Komatsu, H. (2004). Representation of angles embedded within contour stimuli in area V2 of macaque monkeys. *J. Neurosci. 24*(13), 3313–3324.

Karklin, Y. & Lewicki, M. S. (2009). Emergence of complex cell properties by learning to generalize in natural scenes. *457*(7225), 83–86.

Kellner, C. J. & Wachtler, T. (2016). Stimulus size dependence of hue changes induced by chromatic surrounds. *J. Opt. Soc. Am. A*, *33*(3), A267–A272.

Kiper, D. C., Fenstemaker, S. B., & Gegenfurtner, K. R. (1997). Chromatic properties of neurons in macaque area V2. *Vis. Neurosci. 14*(6), 1061–1072.

Klauke, S. & Wachtler, T. (2015). 'Tilt' in color space: Hue changes induced by chromatic surrounds. *J. Vis. 15*(13), 17.

Komatsu, H. (1998). Mechanisms of central color vision. *Curr. Opin. Neurobiol. 8*(4), 503–508.

Komatsu, H., Ideura, Y., Kaji, S., & Yamane, S. (1992). Color selectivity of neurons in the inferior temporal cortex of the awake macaque monkey. *J. Neurosci. 12*(2), 408–424.

Kornblith, S., Cheng, X., Ohayon, S., & Tsao, D. Y. (2013). A Network for Scene Processing in the Macaque Temporal Lobe. *Neuron*, *79*(4), 766–781.

Kuffler, S. W. (1953). Discharge Patterns and Functional Organization of Mammalian Retina. *J. Neurophysiol. 16*(1), 37–68.

Kuriki, I., Sun, P., Ueno, K., Tanaka, K., & Cheng, K. (2015). Hue Selectivity in Human Visual Cortex Revealed by Functional Magnetic Resonance Imaging. *Cereb. Cortex*, *25*(12), 4869–4884.

Laughlin, S. B. (1981). A simple coding procedure enhances a neuron's information capacity. *Z. Naturforsch. C Bio. Sci. 36*(9-10), 910–912.

Laughlin, S. B. & Hardie, R. C. (1978). Common strategies for light adaptation in the peripheral visual systems of fly and dragonfly. *J. Comp. Physiol. A Neuroethol. Sens. Neural. Behav. Physiol. 128*(4), 319–340.

Lee, B. B. (2011). Visual pathways and psychophysical channels in the primate. *J. Physiol. (Lond.) 589*(Pt 1), 41–47.

Lee, B. B. (2014). Color coding in the primate visual pathway: a historical view. *J. Opt. Soc. Am. A*, *31*(4), A103–12.

Lee, B. B., Kremers, J., & Yeh, T. (1998). Receptive fields of primate retinal ganglion cells studied with a novel technique. *Vis. Neurosci. 15*(1), 161–175.

Lee, B. B., Shapley, R. M., Hawken, M. J., & Sun, H. (2012). Spatial distributions of cone inputs to cells of the parvocellular pathway investigated with cone-isolating gratings. *J. Opt. Soc. Am. A*, *29*(2), A223–A232.

Lee, C., Rohrer, W. H., & Sparks, D. L. (1988). Population coding of saccadic eye movements by neurons in the superior colliculus. *Nature*, *332*(6162), 357–360.

Lee, H., Ekanadham, C., & Ng, A. Y. (2008). Sparse deep belief net model for visual area V2. In J. C. Platt, D. Koller, Y. Singer, & S. T. Roweis (Eds.), *NIPS* (pp. 873–880).

Lee, S. C. S. & Grünert, U. (2007). Connections of diffuse bipolar cells in primate retina are biased against S-cones. *J. Comp. Neurol. 502*(1), 126–140.

Lee, T.-W., Wachtler, T., & Sejnowski, T. J. (2002). Color opponency is an efficient representation of spectral properties in natural scenes. *Vis. Res. 42*(17), 2095–2103.

Lennie, P., Haake, W. P., & Williams, D. R. (1991). The Design of Chromatically Opponent Receptive Fields. In M. S. Landy & J. A. Movshon (Eds.), *Computational models of visual processing* (pp. 71–82).

Lennie, P., Krauskopf, J., & Sclar, G. (1990). Chromatic mechanisms in striate cortex of macaque. *J. Neurosci. 10*(2), 649–669.

Lewicki, M. S. (2002). Efficient coding of natural sounds. *Nat. Neurosci. 5*(4), 356–363.

Linsker, R. (1988). Self-organization in a perceptual network. *Computer*, *21*(3), 105–117.

Livingstone, M. S. & Hubel, D. H. (1984). Anatomy and physiology of a color system in the primate visual cortex. *J. Neurosci. 4*(1), 309–356.

Lonini, L., Forestier, S., Teulière, C., Zhao, Y., Shi, B. E., & Triesch, J. (2013). Robust active binocular vision through intrinsically motivated learning. *Front. Neurorobot. 7*.

MacKay, D. J. C. (2003). *Information Theory, Inference and Learning Alorithms*. Cambridge: Cambridge University Press.

MacLeod, D. I. & Boynton, R. M. (1979). Chromaticity diagram showing cone excitation by stimuli of equal luminance. *J. Opt. Soc. Am. A*, *69*(8), 1183–1186.

Maloney, L. T. (1986). Evaluation of linear models of surface spectral reflectance with small numbers of parameters. *J. Opt. Soc. Am. A*, *3*(10), 1673–1683.

Martin, P. R., Lee, B. B., White, A. J., Solomon, S. G., & Rüttiger, L. (2001). Chromatic sensitivity of ganglion cells in the peripheral primate retina. *410*(6831), 933–936.

Michael, C. R. (1978a). Color vision mechanisms in monkey striate cortex: dual-opponent cells with concentric receptive fields. *J. Neurophysiol. 41*(3), 572–588.

Michael, C. R. (1978b). Color-sensitive complex cells in monkey striate cortex. *J. Neurophysiol.*

Mollon, J. D. (2003). The origins of modern color science. In S. K. Shevell (Ed.), *The science of color*. Washington: Elsevier.

Mollon, J. D. (2006). Monge: The Verriest lecture, Lyon, July 2005. *Vis. Neurosci. 23*(3-4), 297–309.

Mollon, J. D. & Jordan, G. (1997). On the nature of unique hues. In *John dalton's colour vision legacy*. London, UK: Taylor & Francis.

Nadal, J.-P. & Parga, N. (1994). Non Linear Neurons in the Low Noise Limit: A Factorial Code Maximizes Information Transfer. *Network*, *5*, 565–581.

Nassi, J. J. & Callaway, E. M. (2009). Parallel processing strategies of the primate visual system. *Nat. Rev. Neurosci. 10*(5), 360–372.

Newton, I. (1671). A Letter of Mr. Isaac Newton, Professor of the Mathematicks in the University of Cambridge; Containing His New Theory about Light and Colors: Sent by the Author to the Publisher from Cambridge, Febr. 6. 1671/72; In Order to be Communicated to the R. Society. *Phil. Trans. R. Soc. Lond. 6*(69-80), 3075–3087.

Newton, I. (1730). *Opticks*. A treatise of the reflections, refractions, inflections and colours of light. London: Printed for William Innys at the West-End of St. Paul's.

Olshausen, B. A. & Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, *381*(6583), 607–609.

Paradiso, M. A. & Nakayama, K. (1991). Brightness perception and filling-in. *Vis. Res. 31*(7-8), 1221–1236.

Paulus, W. & Körger-Paulus, A. (1983). A new concept of retinal colour coding. *Vis. Res. 23*(5), 529–540.

Pinna, B., Brelstaff, G., & Spillmann, L. (2001). Surface color from boundaries: a new 'watercolor' illusion. *Vis. Res. 41*(20), 2669–2676.

Pinsk, M. A., DeSimone, K., Moore, T., Gross, C. G., & Kastner, S. (2005). Representations of faces and body parts in macaque temporal cortex: a functional MRI study. *Proc. Natl. Acad. Sci. U.S.A. 102*(19), 6996–7001.

Pitkow, X. & Meister, M. (2012). Decorrelation and efficient coding by retinal ganglion cells. *Nat. Neurosci. 15*(4), 628–635.

Pouget, A., Beck, J. M., Ma, W. J., & Latham, P. E. (2013). Probabilistic brains: knowns and unknowns. *Nat. Neurosci. 16*(9), 1170–1178.

Pouget, A., Dayan, P., & Zemel, R. (2000). Information processing with population codes. *Nat. Rev. Neurosci. 1*(2), 125–132.

Pouget, A. & Latham, P. E. (2007). Population Codes. In *Bayesian brain* (pp. 893–897). Cambridge, MA: MIT Press.

Quiroga, R. Q., Reddy, L., Kreiman, G., Koch, C., & Fried, I. (2005). Invariant visual representation by single neurons in the human brain. *Nature*, *435*(7045), 1102–1107.

Reid, R. C. & Shapley, R. M. (1992). Spatial structure of cone inputs to receptive fields in primate lateral geniculate nucleus. *Nature*, *356*(6371), 716–718.

Reid, R. C. & Shapley, R. M. (2002). Space and time maps of cone photoreceptor signals in macaque lateral geniculate nucleus. *J. Neurosci. 22*(14), 6158–6175.

Rich, D., Cazettes, F., Wang, Y., Peña, J. L., & Fischer, B. J. (2015). Neural representation of probabilities for Bayesian inference. *J Comput Neurosci*, *38*(2), 315–323.

Ruderman, D. L., Cronin, T. W., & Chiao, C. C. (1998). Statistics of cone responses to natural images: implications for visual coding. *J. Opt. Soc. Am. A*, *15*, 2036–2045.

Rushton, W. A. (1972). Pigments and signals in colour vision. *J. Physiol. (Lond.) 220*(3), 1–31.

Scheibner, H. M. O. & Boynton, R. M. (1968). Residual Red-Green Discrimination in Dichromats. *J. Opt. Soc. Am. 58*(8), 1151–1158.

Schmid, A. M., Purpura, K. P., & Victor, J. D. (2014). Responses to Orientation Discontinuities in V1 and V2: Physiological Dissociations and Functional Implications. *J. Neurosci. 34*(10), 3559–3578.

Schölkopf, B. & Smola, A. J. (2002). *Learning with kernels: support vector machines, regularization, optimization, and beyond*. Cambridge, MA: MIT Press.

Shan, H. & Cottrell, G. (2013). Efficient Visual Coding: From Retina To V2. *arXiv*. arXiv: 1312.6077 [q-bio.NC]

Shannon, C. E. (1948). The mathematical theory of communication. *Bell System Technical Journal*, *27*(3), 379–432.

Sherrington, C. S. (1906). Observations on the scratch-reflex in the spinal dog. *J. Physiol. (Lond.) 34*(1-2), 1–50.

Simoncelli, E. P. (2003). Vision and the statistics of the visual environment. *Curr. Opin. Neurobiol. 13*(2), 144–149.

Simoncelli, E. P. & Olshausen, B. A. (2001). Natural image statistics and neural representation. *Annu. Rev. Neurosci. 24*(1), 1193–1216.

Smith, T. & Guild, J. (1931). The C.I.E. colorimetric standards and their use. *Trans. Opt. Soc. 33*(3), 73–134.

Srinivasan, M. V., Laughlin, S. B., & Dubs, A. (1982). Predictive coding: a fresh view of inhibition in the retina. *Proc. R. Soc. Lond., B, Biol. Sci. 216*(1205), 427–459.

Stockman, A. & Sharpe, L. (2000). The spectral sensitivities of the middle- and long-wavelength-sensitive cones derived from measurements in observers of known genotype. *Vis. Res. 40*(13), 1711–1737.

Svaetichin, G. & Macnichol, E. F. (1959). Retinal mechanisms for chromatic and achromatic vision. *Ann. N. Y. Acad. Sci. 74*(2), 385–404.

Szmajda, B. A., Buzás, P., FitzGibbon, T., & Martin, P. R. (2006). Geniculocortical relay of blue-off signals in the primate visual system. *PNAS*, *103*(51), 19512–19517.

Tailby, C., Szmajda, B. A., Buzás, P., Lee, B. B., & Martin, P. R. (2008). Transmission of blue (S) cone signals through the primate lateral geniculate nucleus. *J. Physiol. (Lond.) 586*(24), 5947–5967.

Tanaka, K. (1996). Inferotemporal Cortex and Object Vision. *Annu. Rev. Neurosci. 19*(1), 109–139.

Tanaka, K. (2003). Columns for complex visual object features in the inferotemporal cortex: clustering of cells with similar but

slightly different stimulus selectivities. *Cereb. Cortex*, *13*(1), 90–99.

Thorell, L. G., De Valois, R. L., & Albrecht, D. G. (1984). Spatial mapping of monkey V1 cells with pure color and luminance stimuli. *Vis. Res. 24*(7), 751–769.

Tokutake, Y. & Freed, M. A. (2008). Retinal ganglion cells - spatial organization of the receptive field reduces temporal redundancy. *Eur. J. Neurosci. 28*(5), 914–923.

Tsao, D. Y., Freiwald, W. A., Tootell, R. B. H., & Livingstone, M. S. (2006). A Cortical Region Consisting Entirely of Face-Selective Cells. *Science*, *311*(5761), 670–674.

Ts'o, D. Y. & Gilbert, C. D. (1988). The organization of chromatic and spatial interactions in the primate striate cortex. *J. Neurosci. 8*(5), 1712–1727.

Valberg, A. (2001). Unique hues: an old problem for a new generation. *Vis. Res. 41*(13), 1645–1657.

Valberg, A., Lee, B. B., & Tigwell, D. A. (1986). Neurones with strong inhibitory s-cone inputs in the macaque lateral geniculate nucleus. *Vis. Res. 26*(7), 1061–1064.

Van Horn, S. C., Erişir, A., & Sherman, S. M. (2000). Relative distribution of synapses in the A-laminae of the lateral geniculate nucleus of the cat. *J. Comp. Neurol. 416*(4), 509–520.

Viénot, F., Brettel, H., Ott, L., Ben M'Barek, A., & Mollon, J. D. (1995). What do colour-blind people see? *Nature*, *376*(6536), 127–128.

Vinje, W. E. & Gallant, J. L. (2000). Sparse coding and decorrelation in primary visual cortex during natural vision. *Science*, *287*(5456), 1273–1276.

Vinje, W. E. & Gallant, J. L. (2002). Natural stimulation of the nonclassical receptive field increases information transmission efficiency in V1. *J. Neurosci. 22*(7), 2904–2915.

Von der Heydt, R. & Pierson, R. (2006). Dissociation of color and figure-ground effects in the watercolor illusion. *Spatial vision*, *19*(2-4), 323–340.

von Helmholtz, H. (1867). *Handbuch der physiologischen Optik*. Hamburg: Voss.

Wachtler, T., Dohrmann, U., & Hertel, R. (2004). Modeling color percepts of dichromats. *Vis. Res. 44*(24), 2843–2855.

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. *J. Vis. 7*(8), 1–14.

Wachtler, T., Lee, T.-W., & Sejnowski, T. J. (2001). Chromatic structure of natural scenes. *J. Opt. Soc. Am. A*, *18*(1), 65–77.

Wachtler, T., Sejnowski, T. J., & Albright, T. D. (2003). Representation of color stimuli in awake macaque primary visual cortex. *Neuron*, *37*(4), 681–691.

Walraven, J. (1973). Spatial characteristics of chromatic induction; The segregation of lateral effects from straylight artefacts. *Vis. Res. 13*(9), 1739–1753.

Webster, M. A., Miyahara, E., Malkoc, G., & Raker, V. E. (2000a). Variations in normal color vision. I. Cone-opponent axes. *J. Opt. Soc. Am. A*, *17*(9), 1535–1544.

Webster, M. A., Miyahara, E., Malkoc, G., & Raker, V. E. (2000b). Variations in normal color vision. II. Unique hues. *J. Opt. Soc. Am. A*, *17*(9), 1545–1555.

Webster, M. A. & Mollon, J. D. (1997). Adaptation and the color statistics of natural images. *Vis. Res. 37*(23), 3283–3298.

Williams, D. R., Sekiguchi, N., Haake, W., Brainard, D., & Packer, O. (1991). The Cost of Trichromacy for Spatial Vision. In *From pigments to perception* (pp. 11–22). Boston, MA: Springer US.

Willmore, B. & Tolhurst, D. J. (2001). Characterizing the sparseness of neural codes. *Network*, *12*(3), 255–270.

Witzel, C. & Gegenfurtner, K. R. (2013). Categorical sensitivity to color differences. *J. Vis. 13*(7), 1–1.

Wuerger, S. M., Atkinson, P., & Cropper, S. (2005). The cone inputs to the unique-hue mechanisms. *Vis. Res. 45*(25-26), 3210–3223.

Young, T. (1802). The Bakerian Lecture: On the Theory of Light and Colours. *Phil. Trans. R. Soc. Lond. 92*, 12–48.

Zweig, S., Zurawel, G., Shapley, R., & Slovin, H. (2015). Representation of Color Surfaces in V1: Edge Enhancement and Unfilled Holes. *J. Neurosci. 35*(35), 12103–12115.

PUBLICATION LIST

ARTICLES

Kellner, C. J., Stoewer, A., Benda, J., Wachtler, T., & Grewe, J. (n.d.). *Acquire together, file together — Integrated Handling of Data and Metadata in the NIX Project*. in preparation.

Kellner, C. J. & Wachtler, T. (n.d.). *Population coding model of color in visual cortex*. in preparation.

Kellner, C. J. & Wachtler, T. (2013). A distributed code for color in natural scenes derived from center-surround filtered cone signals. *Front. Psychol. 4*(661).

Kellner, C. J. & Wachtler, T. (2016). Stimulus size dependence of hue changes induced by chromatic surrounds. *J. Opt. Soc. Am. A*, *33*(3), A267–A272.

Sobolev, A., Stoewer, A., Leonhardt, A., Rautenberg, P. L., Kellner, C. J., Garbers, C., & Wachtler, T. (2014). Integrated platform and API for electrophysiological data. *Front. Neuroinform. 8*, 32.

Sobolev, A., Stoewer, A., Pereira, M., Kellner, C. J., Garbers, C., Rautenberg, P. L., & Wachtler, T. (2014). Data management routines for reproducible research using the G-Node Python Client library. *Front. Neuroinform. 8*.

Sommer, F., Wachtler, T., Teeters, J., Benda, J., Davison, A., Eglen, S., … Wark, B. (2016). Requirements for storing electrophysiology data. arXiv: 1605.07673

SELECTED CONFERENCE PRESENTATIONS

Felmy, F., Berner, S., Franzen, D., Blank, S., & Kellner, C. (2013). Calcium entry sites of neurons in the medial superior olive. In *Göttingen meeting of the german neuroscience society*.

Garbers, C., Kellner, C., Mokrzycki, L., & Wachtler, T. (2012). Neuronal data storage using document oriented databases. In *Conference abstract: bernstein conference* (264). doi:10.3389/conf.fncom.2012.55.00264

Haas, O., Kellner, C., & Wachtler, T. (2013). A model of lateral interactions in color vision. In *Göttingen meeting of the german neuroscience society*.

Kellner, C. J., Stoewer, A., Sonntag, M., Sobolev, A., Benda, J., Wachtler, T., & Grewe, J. (2016). Consistent data organization made easy: a versatile file format for neuroscience. In *4th incf japan node international workshop advances in neuroinformatics (aini) 2016 and14th incf nodes workshop*.

Kellner, C. J. & Wachtler, T. (2014). Lateral interactions in an anisotropic population code for color predict human color induction effects. In *2014 neuroscience meeting planner*. Washington, DC: Society for Neuroscience.

Kellner, C. J. & Wachtler, T. (2015). Stimulus size dependence of hue changes induced by chromatic surrounds. In *The 23rd symposium of the international colour vision society*.

Kellner, C. & Wachtler, T. (2012a). Accessing electrophysiological data from python. In *Conference abstract: 5th incf congress of neuroinformatics* (42). doi:10.3389/conf.fninf.2014.08.00042

Kellner, C. & Wachtler, T. (2012b). Nonlinear processing of natural stimuli explains multiple color categories in dichromats. In *Conference abstract: bernstein conference* (104). doi:10.3389/conf.fncom.2012.55.00104

Wachtler, T., Sobolev, A., Kellner, C. J., Le Franc, Y., & Grewe, J. (2014). Neuroinformatics for efficient data management and reproducibility in electrophysiology. In *2014 neuroscience meeting planner*. Washington, DC: Society for Neuroscience.

ACKNOWLEDGEMENTS

So long and thanks for all the fish

— Douglas Adams

I have had the great joy and pleasure to work on this thesis during the last few years, and it has been quite a ride. I benefited from great moral support, with knowledge and advice from many people. It would not have been possible to do this without them — and it certainly would have been less fun! First and foremost I want to thank Thomas Wachtler, my supervisor, who gave a Philosopher and Computer Scientist the chance to dive into the fascinating field of neuroscience. Thomas always provided me with so much freedom and independence to do research, programming and learn new techniques on my own. But at the same time he was always supportive and encouraging, and often gave me the little nudge in the right direction that I needed.

I also want to thank my thesis advisory committee, Stefan Glasauer and Andreas Herz, who provided help and critical comments. The Graduate School of Systemic Neurosciences must not be forgotten, with the nice people of the administration, who were always helpful; also the other students, who provided such a pleasant environment in which to discuss ideas.

I need to thank Satra Gosh for hosting me at the MIT, an experience that provided me with many new insights and ideas. Also Heidi and Mark, together with Kevin and Bernard for their warm welcome in Boston. Of course, many thanks have to be directed to the rest of the Wachtler Lab, and in particular Christian Leibig and Christian Garbers, for many excellent philosophical and political discussions. Not to be missed are the members of the G-Node, Adrian Stoewer, Michael Sonntag, Andrey Sobolev and Philipp Rautenberg for many interesting and informative discussions on software architecture and programming languages. Also Jan Grewe, who was basically

part of the team, always helping with programming and discussions, and who became such a good friend. I am also very thankful to my former mentor, Thomas Mohrs — without his guidance and friendship I would not have become the person that I am.

Finally, I want to thank all the great friends that made the last years worthwhile, especially Sarah and Franzi who not only helped me survive the drudgery of *the practical courses*, but also made them fun; also Olivia and Laurence for generally being awesome friends. My biggest thanks go to my parents for their enduring support and, of course, most importantly to Delwen for her love and never ending patience.

EIDESSTATTLICHE VERSICHERUNG/AFFIDAVIT

Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation *Coding Mechanisms & Contextual Interactions in Color Vision* selbstständig angefertigt habe, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

I hereby confirm that the dissertation *Coding Mechanisms & Contextual Interactions in Color Vision* is the result of my own work and that I have only used sources or materials listed and specified in the dissertation.

München, im 5th July 2016
Munich, 5th July 2016

Christian J. Kellner

AUTHOR CONTRIBUTIONS

The contributions of the authors Christian J. Kellner (CJK) and Thomas Wachtler (TW) to the studies included in this cumulative thesis are as follows:

Kellner, C. J. & Wachtler, T. (2013). **A distributed code for color in natural scenes derived from center-surround filtered cone signals**. *Front. Psychol. 4*(661). doi: 10.3389/fpsyg.2013.00661. CJK and TW designed the study. CJK implemented and performed the simulations. CJK and TW wrote and revised the manuscript.

Kellner, C. J. & Wachtler, T. (2016). **Stimulus size dependence of hue changes induced by chromatic surrounds**. *J. Opt. Soc. Am. A*. *33*(3), A267–A272. doi: 10.1364/JOSAA.33.00A267. CJK and TW designed the study. CJK implemented and supervised the experiment, and also analyzed the data. CJK and TW discussed the data and results. CJK wrote and revised the manuscript with the help of TW.

Kellner C. J. & Wachtler T. (n.d.). **Population coding model of color in visual cortex**. CJK and TW designed the study. CJK implemented and performed the simulation. CJK drafted the manuscript included in this thesis.

We assert that aforementioned author contributions are correct and accurate:

Christian J. Kellner

PD Dr. Thomas Wachtler