Contents

1
Retinal processing and its used for digital imaging

Alleysson David and Guyader Nathalie

To perceive the surrounding world, its beauty and subtlety, we not only move eyes but our visual system makes an incredible job. Even if vision is the most studied sense it is still not clear how the neuronal circuitry of our visual system build a coherent percept of the world. In this chapter, we introduce some pathways between our knowledge of the retinal biology and computational models that might be used in computer vision systems. Because such pathway must encompass several different disciplines, we will certainly be partial in our description. But, this introduction brings all the things that should be taken together to be able to understand the problematic of computer vision.

The first part of the chapter is a short description of the anatomy and physiology of the retina. Useful link will be given for the reader who wants to have deeper explanations. The second part summarizes some models of the retina and its visual functions. We draw a large panorama of these models and we compare them in their performance and usefulness both to explain vision and to be used use in computer vision. The last part is dedicated to some examples of the utilization of models of vision in the context of digital color camera processing. This last part allows us to illustrate our purpose through real applications.

1.1
Introduction

Biologically-inspired computer vision is not solely based on the scientific results on the field of biology itself and psychophysics experiments bring needed data for example to test the efficiency of models. The contrast sensitivity function (CSF) that represents the perceived luminance contrast as a function of the spatial frequencies has been measured before the discovery of the neural network underlying it. Similarly, trichromacy (the fact that color vision lie on a tri-dimensional space) have been stated before we discover the three kinds of cones in the retina. Currently, in most cases, psychophysics precedes biology. In some sense, psychophysics remains biologically-inspired because everybody agreed that human visual behaviors are rooted and emerges with the neural circuitry of the visual system. Often, it is simpler to find a model of human behavior instead of biological structure and function

of the neurons that support it. With the main goal to describe biologically-inspired computer vision, this chapter mixes results that come from psychophysics to results that come from biology for the goal of biologically-inspired computer vision.

The relationship between behavior and underlying biology is still not fully understood. For example, it is still unclear why our visual world appears stable when our eyes move continuously. Especially the neurons and mechanism involve in this processing are not known. Another question arises to explain the contrast sensitivity function: at which level this processing occurs: in the retina, in the transmission between retina and cortex or in cortical areas? In fact, to know exactly at which level this phenomenon is coded is not essential for models of vision. In fact their aim is to reproduce functionalities of human vision to develop signal processing algorithms that mimic the visual system functions. For this, the limit of the imitation of nature by artificial element is not a barrier, if we are able to reproduce behaviors. However, we must careful that a model, even perfectly functioning for behavior, is not necessarily what exactly happen in the visual system.

1.2
Anatomy and physiology of the retina cells

1.2.1
Overview of the retina

Because vision is the most studied sense, a huge number of studies in biology, physiology, medicine, neuroanatomy, psychophysics and also computational modeling, provide data to better understand this sense. Vision starts with the retina, in the back of the eyes, that transmits the light wavelengths and converts it. This chapter goes through the basis of the retina functioning rather than detailed description. The aim is to help a "vision researcher" to roughly understand the functioning of the retina and how to model it for applications. If one wants to know more about the retina anatomy and physiology he/she might want to read two free online books: (1) Webvision, The Organization of the Retina and Visual System by Helga Kolb and colleagues (http://webvision.med.utah.edu/) and (2) Eye, brain and vision by David Hubel (http://hubel.med.harvard.edu/book/bcontex.htm).

The first detailed description of the retina anatomy was provided by Ramon y Cajal one century ago. Since, with the technology and technics improvements, detailed knowledge about the organization of the retina and the visual system was achieved. A schematic representation of the retina is displayed in the Figure 1.1. This schema is a simplified representation because the retina is much more complex and contains more cell types. Basically, the retina layers are organized into columns of cells from the "input" cells, the photoreceptors, rods and cones, to the "output" cells, the retinal ganglion cells that have their synapses directly connected to the primary visual cortex. Layers of cells are inversed compare to the transmission of light, with the photoreceptors located at the back of the retina. Two main layers of cells are generally described in the literature: (1) the outer plexiform layer (OPL), with the photorecep-
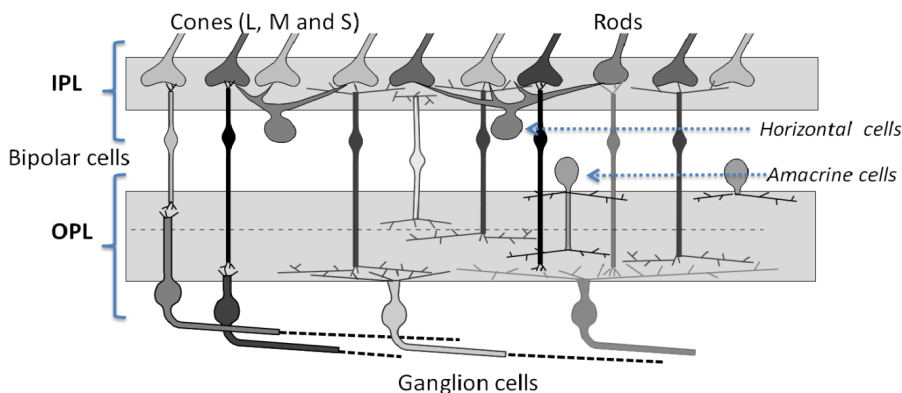
Figure 1.1 Schematic representation of the organization into columns of retinal cells. Light goes through the retina from photoreceptors, located at the inner plexiform layer (IPL), to ganglion cells located in the outer plexiform layer (OPL), passing through bipolar cells (adapted from Hérault, 2010 [43]).

tors, horizontal and bipolar cells, (2) the inner plexiform layer (IPL), with the bipolar, amacrine and ganglion cells.

The photoreceptors convert the light, i.e. incoming photons, into an electrical signal through membrane potential; this operation is called transduction, a phenomenon that exists for all the sensory cells. The electrical signal is transmitted to bipolar and ganglion cells. The electrical signal is integrated within ganglion cells and transformed into nerve spikes. Nerve spikes are a time-coded digital form of an electrical signal. It is used to transmit nervous system information over long distances, in this case through the optic nerve and into the brain visual cortical areas.

1.2.2
Photoreceptor

The retina contains an area without any photoreceptor where ganglion axons connect the retina to the primary visual cortex (back of the brain, occipital lobe) forming the optic nerve. The center of the retina called the fovea (corresponding to the optic axis) is a particular area of the retina with the highest density of photoreceptors; it is located closed to the retina center and corresponds to the central vision. The area of the visual field that is gazed, directly projected onto this central area. There are two main types of photoreceptors, which are named according to their physical shapes: the rods and the cones. The fovea only contains cones (Figure 1.2), which are the photoreceptors sensitive to color information. Three types of cones might be distinguished for their sensitivity to three different wavelengths: S-cones sensitive to S-hort wavelength that responds to bluish color, M-cones sensitive to M-edium wavelength that corresponds to greenish-yellowish and L-cones sensitive to L-ong wavelength that corresponds to reddish. These photoreceptors are activated during day-vision (high light inten-
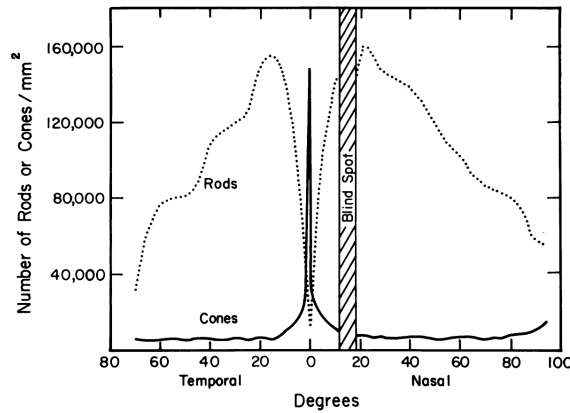
Figure 1.2 Number of photoreceptors (rods and cones) at the retina surface (in square millimeters) as a function of the eccentricity (distance from the fovea center, located at 0). The blind spot corresponds to the optic nerve (an area without photoreceptor). From [62]

sity). The number of cones drastically decreases with the eccentricity onto the retina surface. On the fovea periphery the number of cones is very small compared to the number of rods. These later photoreceptors are responsible for the scotopic vision (vision under low light conditions).

The arrangement of photoreceptors on the retina surface has some consequences on our visual perception. Our visual acuity is maximal for the region of the visual field gazed at (central vision) whereas it decreases rapidly for the surrounding areas (peripheral vision). It is important to note that the decrease of our visual acuity in peripheral vision is only partly due to the retina photoreceptor repartition. In fact other physiological and anatomical explanations might be given (see 1.2.3). Moreover, whereas we do not feel it in our everyday life only our central vision is colored, peripheral vision is achromatic. Another particularity of cones arrangement is that the juxtaposition of L, M and S cones is random in the retina (See Figure 1.4(b)) and differ from individual to individual.

Photoreceptors dynamically adapt their responses to various levels of illumination. This property allow human vision to be functioning for several decades of light intensity. In fact we perceive the world and we are able to recognize a scene under sunny day or in the twilight.

1.2.3
Outer and Inner plexiform layers (OPL and IPL)

At this stage, it is important to introduce the notion of *receptive field*. The receptive field of a cell is necessary to explain the functioning of retinal cells, and more generally, the visual cells. This notion was introduced by Hartline for retinal ganglion

cells [41]. According to the definition of Hartline the receptive field of retinal ganglion cell corresponds to a restricted region of the visual field where light elicited an electrical response of the cell. This notion was extended to other retinal cells as well as cortical visual and sensory cells. The receptive fields of visual cells corresponds to a small volume of the visual field. For example, the receptive field of a photoreceptor is a cone-shaped volume with the different directions in which light elicits a response from the photoreceptor [8]. However, most of the time the receptive field of a visual cell is schematized in two dimensions corresponding to a visual stimulus (a digital image) located at a fixed distance from the participant.

The OPL corresponds to the connection between photoreceptors, bipolar and horizontal cells. This connection is responsible for the center/surround receptive fields displayed by bipolar cells. Two types of bipolar cells might be distinguished: *on-center* and *off-center*. The receptive fields of retinal bipolar cells correspond to two concentric circles with different luminance contrast polarities. On-center cells respond when a light spot falls inside the central circle surrounding by a dark background and off-center cells respond for a dark stimulus surrounding by light [68].The functionality of this antagonist center/surround receptive fields is to enhance contours, as it could be done on an image with a spatial high pass filter.

The IPL corresponds to the connection between bipolar, ganglion and amacrine cells. Ganglion cells still have center/surround receptive fields.

There exist several different retinal ganglion cells that differ in terms of their size and connections. These cells are directly connected to the primary visual cortex via the lateral geniculate nuclei. Retinal ganglion cells vary significantly in terms of their size, connections, and responses. As the number of photoreceptors decreases with the retinal eccentricity, the size of the receptive fields of bipolar and ganglion cells increases with the distance from the fixation point. At the retina center there is the so-called one-to-one connection where one ganglion cell connects to one bipolar cell and to one photoreceptor guarantying maximal resolution around optical axis.Whereas at large eccentricity one bipolar cell connects up to several millions of photoreceptors. Moreover, the size of the receptive fields of bipolar and ganglion cells increases with their eccentricity and at successive processing stages in the visual pathway.

There are essentially three types of ganglion cells, distinct anatomically and functionally, that project differentially in the lateral geniculate nucleus (LGN) [29]. The more numerous (around 80%) in the fovea and near fovea is the midget ganglion cell, called because of their small sizes. They are the cells that connect in one-to-one with photoreceptors. There responses are sustained, slow and they are known to carry high spatial information at low contrast as well as L-M color opponency. They project their axons to the parvocellular part of the LGN. In opposite, parasol ganglion cells are large and connect several bipolars and several photoreceptors even in the fovea. There response are transients and fast, they are known to carry high contrast information at low spatial frequency, they are color blind. They also participate to the rod circuit at night. They project their axons in the magnocellular pathway of the LGN. A third kind called bistratified ganglion cell have their dendrites forming two layers in the IPL, they are known to carry S-(L+M) color opponency and project in the koniocellular pathway of the LGN. Finally, it has been discovered recently a

new kind of ganglion cell containing melanopsin pigment that make them directly sensitive to light.

1.2.4
Summary of the main preprocessing done by the retinal cells

The retinal ganglion cell axons are directly connected to the superior colliculi and the cortical visual areas passing through the lateral geniculate nuclei.

The role of the retina is to pre-process the visual signal for the brain. Then the signal is successively processed by several cortical areas to bring a coherent percept of the world with the recognition of the objects that compose the scene.

The main functions of the retina are:

- to convert the photons into firing rate that can be transmitted between neurons
- to make adaptation to the mean luminance of the scene but also to the local luminance contrast
- to decompose the input signal: low spatial frequencies are transmitted first to the cortex, immediately followed by high spatial frequencies and opponent color information.

1.3
Overview of retina models

Most of the retina models have been developed because our visual system is the most powerful system for object recognition. The aim of retina models is to pre-process the input signal into an efficient representation of the natural scene. Consequently, most of the retina models are based on the so-called ecological vision because the pre-processing of the retina optimizes the representation of information coming from the visual scene. Interesting lights in this context is those who are related to animal evolution, such as the light reflected by natural objects such as fruit, trees, landscape, etc...

Hence, models of the acquisition and pre-process done by the retina mainly focus on the modeling of the retinal cells that encode visual information. More specifically those models address the following question: Are we able to reproduce the measured response of some retinal cells using simple identified elements of the biology. One of the fundamental model for that is the *Hugkins and Huxley's model* [46] which explains the dynamics of neural conduction, relating the conductance along neurons with an electrical equivalent model.

The advantage of defining an equivalent electrical model of the network of simple neural components is double. First, analog signal and image processing could be modeled as an electrical system that can be translated in terms of electrical equations ; in this way a direct implementation based on analog component is easy with a conversion into digital approximation. Second, it is easy to infer the behavior of a heterogeneous network of cells with different functioning and hence different mod-

els to estimate their functioning. With a model of the synapse transmission between neurons, the integration of heterogeneous neural network becomes plausible. This approach have been very fruitful in what is now called, computational models of vision [52, 42]. In the following section, we briefly describe some of models of the retina based on the Hugkins and Huxley's model of neural transmission.

Then, we discuss a more general approach (system approach) of neural modeling; such models build a geometric space of neural activities, and then, infer the behavior of the network using the geometric rules stated at the neural scale. The neurogeometry, term initially used by Jean Petitot is a very seducing example of such model category. Focusing on color vision, the Riemannian geometry has extensively been used for explaining color vision from the LMS cone excitation space. We describe how those *line-element* models have tried to put human color discrimination capabilities in a context of Riemannian geometry. We finish this part about retina models with the information theory model that considers retina as an information transmission system.

1.3.1
Biological models of the retina

1.3.1.1   Cellular and molecular model - Van Hateren

Van Hateren choose cellular and molecular scales to defined his retinal model [77]. His model of the outer retina have been validated against the measured signal on the horizontal cells to wide field spectrally white stimuli. Horizontal cells are known to change their sensitivities and control their bandwidth with the background intensity. The response is non-linear and the non-linear function changes for background ranging from 1-1000 td. All these properties are taken into account in the van Hateren's model and explained by three cascaded feedback loops representing successively the calcium concentration at the cone outer pigment, the membrane voltage of the cone inner pigment and the signal modulation of the horizontal cells to cone signals.

His model shows that cone responses are the major factor regulating sensitivity in the retina. In a second version, they propose to include the pigment bleaching at the cone outer segment to better simulate cone response [80].

Van Hateren implement his model using autoregressive moving-average filter that allow modifying in real time the processing following the adaptation parameter (or the state at time t of the static non linearity part of the model) and fast calculus [77, 80]. He also manage to give all the constants involved in the model and implementation in Fortran and Matlab which make the model very useful. The model is principally adapted to simulating cone responses, but can be used in digital image processing such as high dynamic range processing [79]. As far as we know, models developed by van Hateren are the only molecular model of retina that could be applied on images or image sequences.

Even focusing on the cellular and molecular modeling, van Hateren fall on a more general kind of modeling that does not necessary follow closely biology. These general models use two kinds of building bricks, a linear (L) multi-components filtering

and a static non-linearity (N). The blocks are then considered in cascade in what is called LNL or NLN models to consider either the case of two linear operator *sandwiched* the non-linear one, or the reverse. These kind of models are decomposition of the general non-linear Voltera system [19, 81] because the non-linearity is static and point-wise and the only mixing of components (either spatial or temporal) is through a linear function. These models have been found to be representative of the mechanism of gain control in the retina [71, 2]. The NLN model (also called Hammerstein-Wiener) is less tractable than the LNL model [28, 21] but may potentially be more interesting for modeling vision and specially color vision [2, 84, 87, 25].

Cellular and molecular network models of the retina are the most promising because they are at a perfect scale for fitting electro-physiological measurement and theory of the dynamic of heterogeneous neural network. But they are difficultly tractable because they reach the limit of our knowledge about optimization of complex systems.

1.3.1.2 **Network models - Herault, Worher**

Models that simulate the retina functioning are numerous from detailed models of a specific physiological phenomenon to models of the whole retina. We wanted to emphasize in this chapter on two interesting models that take into account the different aspects of the retina functioning. These models naturally make simplifications regarding the complexity of the retina cells functioning but provide interesting results by replicating experimental results either on particular response cells (for example, the output retina cells of the cats) or on more global perceptual phenomenon (the contrast sensitivity function). An implementation of these two models is provided through executable softwares that can be freely downloaded:

- Virtual retina from Adrien Wohrer & Pierre Kornprobst [86] :
  http://www-sop.inria.fr/odyssee/software/virtualretina/
- The retina model originally proposed by Hérault [44] and improved and implemented by Benoit Alexandre [20]
  https://sites.google.com/site/benoitalexandrevision/download

The model proposed by J. Hérault in his book [43] summarizes the main functions of the retina and it is compared to an image pre-processor. His simplified model of the retinal circuits is based on the electrical properties of the cellular membrane that provides a linear model of the retinal cells. Beaudot and colleagues first proposed a basic electrical model equivalent to a serie of cells connected by gap junctions [16, 17] electrical model This linear model provides the fundamental equations of the retina functions to finally model the retina circuitry through a spatio-temporal filter (see figure 1.3 for its transfer function).

This model provides interesting results first reproducing biological data provided by Buser and Imbert [24]. This model reproduces the main function of the retina with a luminance adaptation, as well as a luminance contrast one that results in a contrast enhancement. Finally due to the non-separability of the spatial and temporal dimension, first the low spatial frequencies of the signal are transmitted and latter refined by the high spatial frequencies that arrive later. This phenomenon reproduces the Coarse to fine processing that occurs in the visual system. The retina model has
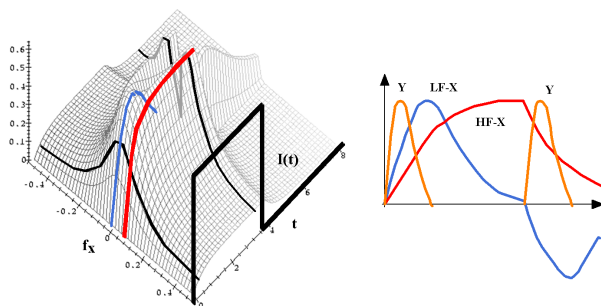
Figure 1.3 (left) The spatio (only in one dimension fx) temporal transfer function of the retina for a square input I(t). (right) The temporal profiles of high spatial frequencies (HF-X) that corresponds to the parvocellular pathway, low spatial frequencies (LF-X) that corresponds to the magnocellular pathway X and Y cells (adapted with permission from [43]

been integrated into a real-time C/C++ optimized program [20]. This paper is more general and proposed with the retina model a model of the cortical cells to show the advantages of bio-inspired models to develop efficient modules for low level image processing.

Wohrer and Kornprobst [86] also propose a bio-inspired model of the retina and propose a simulation software, called Virtual Retina. The proposed model transforms a video into spike trains with two main goals: a large scale simulation (up to 100,000 neurons) in reasonable processing times and a strong biologically plausibility. Their model includes a linear model of filtering in the Outer Plexiform Layer, a shunting feedback at the level of bipolar cells accounting for rapid contrast gain control, and a spike generation process modeling ganglion cells. Their model reproduces several experimental measurements from single ganglion cells such as cat X and Y cells (like the data from Enroth-Cugell and Robson [33]).

1.3.1.3 **Parallel and descriptive model**

To conclude with biological model of retina, we introduce the work of Werblin & Roska [85] based on the identification of severals different processing by cells in the retina. Characteristics of cells have been invested by electro-physiology. They considered then the retina as a highly parallel processing system which provide several maps of information build by the network of different cells. Advantage of this approach is to provide images or images sequences of the different functions of the retina.

Similarly, Gollisch & Meister [39] have recently proposed a paper where they discuss the possibility that the retina would be embedded with high level functions. They show elegantly how these functions could be implemented by the retinal network.

No one knows at which degree of sophistication the retina is build for, but what we know is that almost 98% of the ganglion cell having their dendrite leaving the retina to the LGN have been identified robustly. Nevertheless, their role in our everyday

vision and perception is not yet been understood.

1.3.2
Geometry models

Geometry models have been extensively developed for color vision and object's location in 3D space. The reason is certainly because Riemann [67] argues himself that these two aspects of vision need a more general space than the Euclidean space. For him, they are the only two examples of continuous manifold in three dimensions that cannot be well represented in Euclidean geometry. Apparently, this motivate his research on what is called now, Riemannian geometry. The major distinction between Euclidean and Riemannian geometry is, in the later, that the notion of distance or norm (length of a segment defined by two points) is local and depend from point to point in the space. In contrary in Euclidean geometry norm is fixed.

Soon after Riemann, Helmholtz publishes the first Riemannian model [82] for color discrimination. Before explaining this model in detail, let considers the evolution of psychophysics at that time. It was shown by Weber that a variation of the sensation of weight depends on the weight considered. This law is known as $\frac{\Delta I}{I} = Cst$. Intuitively speaking, if you have a weight in your hand let say, I kilogram, then you feel a difference in weight if you add another weight such that the difference in weight ΔI exceed a constant multiplying I. The same kind of law have been found for light by Bouguer [22]. Formalizing the relationship between the physical space and the psychological space, Fechner [34] proposes that the law of sensation S is given by $S = k*log(I)$ which is one possible solution of the Weber law because the derivative of the Fechner'law gives $\Delta S = \frac{k}{I}\Delta I = Cst$. It should be noticed that Stevens [73] proposes that the law of sensation is given by a power law, $S = k*I^{a}$, considering that a log function also apply for the space of sensation.

The idea of Helmholtz was to provide a model of color discrimination based on the Fechner's law embedded in a three dimensional Riemannian space representing color vision space. The length of element of distance ds is therefore given by:

$$ds^2 = (dlogR)^2 + (dlogG)^2 + (dlogB)^2 = \left(\frac{dR}{R}\right)^2 + \left(\frac{dG}{G}\right)^2 + \left(\frac{dB}{B}\right)^2 \quad (1.1)$$

It is in some sense a direct extension of the Fechner's law in three dimensional space [3]. It is also a Riemannian space because the norm depend on the location of the space, given by the value of R, G and B denominators in the equation. And, locally (i.e. for a fixed R, G and B) the norm is the square norm. This class of discrimination model have been called line-element. Many extension of the Helmholtz's model have been given (see for example Schrödinger [70], Vos and Valraven [83], Stiles [74], Koenderink [48], Alleysson and Hérault [2]).

Despite these theoretical models, the prediction of color discrimination was not so accurate. MacAdam proposes in 1942 a direct measure of the color difference in CIExy space for a fixed luminance. His well known MacAdam's ellipses served as a basis measure for discrimination and the CIE had proposed two non linear space CIE-LUV, CIE-Lab for taking into account those discrimination curves. It has been shown

that MacAdam ellipses could be modeled on a Riemannian space [49] although the corresponding curvature become negative and positive following the position on the color space.

Recently, geometry have been used in a more general context than color metric. It has been used for modeling the processing of assembly in neuron in the visual system in a field called neurogeometry [63, 26]. Using geometry rules based on the orientation function of the map of neuron modeling primary visual cortex V1, Jean Petitot was able to show the spatial link of neural activities that elicits Kaniza's figure illusion. The approach of neurogeometry is very elegant and fruitful and will probably generalized to the whole brain. But; it is still challenging when considering color vision and the sampling of color information with retinal cone mosaic that differs in arrangement and proportion from individual to individual [3].

1.3.3
Information models

Describing a scene viewed by a human observer need to consider every point of the scene as reflecting light. For a complex scene, this light field represents a huge amount of information. It is therefore not possible for a brain to encode every individual scenes that way for enable useful information retrieval. From a biological perspective, the retina has hundred millions of photoreceptors but the optic nerve that conduct information from the retina to the brain comprise solely ten millions of fibers. Those arguments have lead to the idea that a compression of information occurs in the visual system[14, 36].

From a computational point of view, this compression is similar to the compression of sound in mp3 or image in jpeg file format. A transform of the input data is done over a basis that reduce redundancy (the choice of the basis is given by statistical analysis of the input data). The representation of information onto this basis allow reducing the size of the representation.

From a biological point of view, Olshausen and Field [60, 61], found that learning a sparse, redundancy reduced, code from a set of natural images gives basis function that resemble the receptive fields of neuron in the visual system V1. This work have been further extend [18, 9], specifically in the temporal [78, 31] and color domain with the correspondence between opponent color coding by neurons and statistical analysis of natural colored images [23, 69, 75, 54]. A particular treatment of the theory had focused on retinal processing [12, 11, 10].

Ecology of vision, by stating that the visual system is adapted to the natural visual environment, give a simple rule and efficient methods for testing and implementing the principle. However, until now, the network of neuron that could implement such processing is not found. Actually, information theory is a black box, describing well the relationship between what enter into the box and what exit from the box. But, it does not give any rule of what is inside the box, the processing by neurons of visual information.

We have already made other criticisms about redundancy reduction for color vision in the context of sampling color by the cone mosaic in the retina [4, 5]. The idea
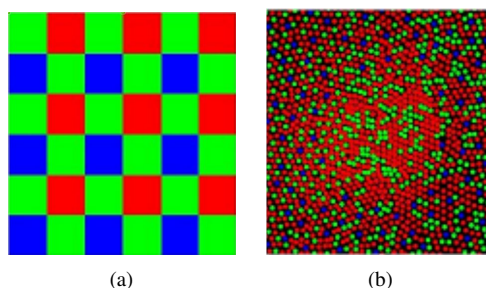
Figure 1.4 (a) Bayer color filter array (CFA) (b) Simulation of the arrangement of L-cone, M-cone and S-cone in a human retina

is based on the fact that the statistical analysis performed in redundancy reduction implicitly suppose that the image is stationary in space (statistic is done over patches of the image and should be coherent from one patch to another). But, because of the random arrangement of cones in the retina mosaic, even if the natural scene is stationary, the sampled signal will not be.

Also, redundancy reduction models seek for a unique model of information representation. But because cone mosaic differ from individual to individual, it is likely that the representation change from individual to individual [1].

1.4
Application to digital photography

The main application of retina models is given in the processing of raw images into digital cameras. In digital cameras, the image of the scene is captured by a sensor usually covered with a matrix of color filters to allow for color acquisition. At each pixel of the sensor a particular filter is attached. The main color filter array (CFA) used in today digital camera is called Bayer CFA from the name of its inventor [15]. This filter array is composed by three colors (either RGB ou CMY) and is a periodic replication of two of the color on one line over two and two other colors in remaining lines. By consequent, one of the three colors is represented twice in the array (usually 1R, 2G and 1B), see Figure 1.4(a). For comparison, Figure 1.4(b) represent a simulation of the random arrangement of L, M and S-cones in the human retina.

At the output of the sensor the image is digitalized into height, twelve or more bits. But, the image is not already a color image because (1) there is only a single color per pixel, (2) color space is camera based, given by the spectral sensitivity of the camera (3) white may not be white following the tint of the light source (4) tones could appear unnatural because of non linearity in the processing of photometric data. This is the purpose of the digital processing inside the camera to convert raw image coming from the sensor to nicely viewed color image. See those papers and books for a review [64, 55, 57].

Classically, the image processing pipeline is composed on the sequential operation. Image acquisition by the sensor, some corrections on raw data (dead pixels, blooming, white pixels, ...), demosaicing for retrieving three color information per pixels, white balancing and tone mapping. In the following section we will described these operations and give reference to literature. We get a parti-pris to not being exhaustive in the description but to focus on study that use directly or indirectly models of vision to state their methods.

1.4.1
Demosaicing

Demosaicing operation is an inverse problem [66] because the operation seek to retrieve intensity in color that are not been sampled by the sensor. At the origin, demosaicing have been invested as an interpolation problem using copy of neighbors pixels to fill-in missing pixels or bilinear interpolation which are easily done by analog electronic element such as delay, adder, etc... But, the emergence of digitalization of images at the output of the sensor has allowed more complex demosaicing.

The specificity of color image with inter-channel correlation have been taken into account and almost all method proposed to interpolate color difference instead of color channel directly [45, 40] for taking account that color channels are correlated. Another specificity of images have been take into account, the fact that image are two dimensional, composed by flat surface as well as contours. By identifying the type of area present in a local neighborhood, it is possible to interpolate along edges rather than across them. This property of color image have been extensively used for demosaicing [47, 45]. Conjointly some correction methods have been proposed to improve the quality of image have been first demosaiced [47, 45, 40].

Another approach developed originally by Alleysson et al. [6, 7] and Dubois [32] consists of modeling the expectation of the spatial frequency spectrum of an image acquired through a mosaic. In the case of the Bayer CFA of Figure 1.4, the spatial frequency spectrum shows nine zones of energy, one in the center correspond to luminance (or achromatic information) whereas the height in the border of the spectrum correspond to chrominance (or color information at constant luminance). See Figure 1.5 for an illustration. The demosaicing consist then in estimating (adaptively to image content or not) luminance and chrominance components by selecting in the spectrum information corresponding either to luminance or chrominance.

This approach of demosaicing have been shown to be very fruitful. First it is totally linear method which could easily be optimized for a particular application or for any CFA. Lu *et al.* [56] have shown the theoretical equivalence of frequency selection with alternating projection on convex sets (sets given by pixels and spatial frequency elements of the image) proposed by Gunturk *et al.* [40] whereas Condat [27] shows the equivalence with variational approach. This model of information representation in a CFA image is also strongly related to computational model of vision, but include color component. Nevertheless, the random nature of the cone mosaic in the human retina and the difference in arrangement from individual to individual make the analogy difficult to establish formally. The role of eye movement in the reconstruction of
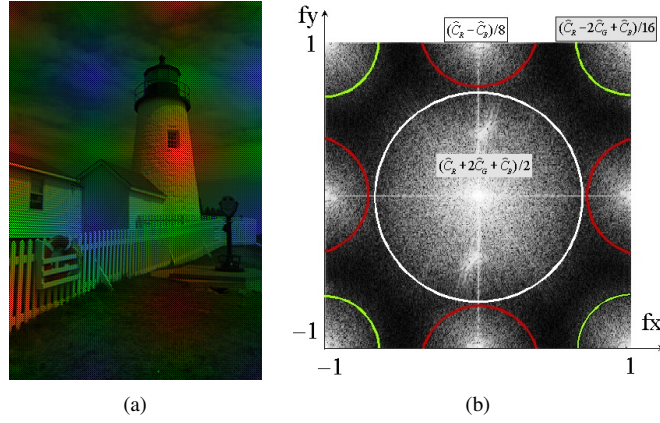
Figure 1.5 (a) Color image acquired through the Bayer CFA (b) The corresponding spatial frequency spectrum showing luminance (R+2G+B) information in the center and chromatic information (R-B, R-2G+B) in the borders.

continuous visual information inside the brain is still to be discovered.

1.4.2
Color constancy - chromatic adaptation - white balance - tone mapping

Retina and visual system show a large extend of adaptation [72] allowing human vision to be accurate even if the dynamic range of the scene is very deep. Also, viewed color of objects are very stable despite of a large difference in their reflective light due to change in illumination (think of the spectrum of a banana under daylight or indoor incandescent light which are quite different, but the banana still appear yellow under both lighting). The so-called color constancy phenomena is a consequence of chromatic adaptation that happen in the retina and visual system as illustrated by Land's retinex theory [51]. A large amount of literature have been produced on retinex theory and its application to image processing.

The principal effect on image quality of not implementing chromatic adaptation in a digital camera could be seen because white on objects or surface do not longer appear white.

From a computational point of view, chromatic adaptation is often based on the von Kries model which stipulate that adaptation is a gain control in the responses of LMS cones. Roughly, von Kries model of chromatic adaptation write $\mathbf{y} = \mathbf{MDM'x}$. Where $\mathbf{x}$ is the vector containing pre-adapted image, $\mathbf{y}$ is the vector containing post-adapted image, $\mathbf{M}$ is the 3x3 unitary matrix ($\mathbf{M'} = \mathbf{M}^{-1}$) for the transformation from image in camera color space to image in LMS color space and $\mathbf{D}$ is the adaptation

transform given by:

$$\mathbf{D} = \left[\begin{array}{ccc} L_2/L_1 & 0 & 0 \\ 0 & M_2/M_1 & 0 \\ 0 & 0 & S_2/S_1 \end{array} \right] \tag{1.2}$$

where $._1$ represent adaptation under illuminant 1 and $._2$ adpatation under illuminant 2.

By studying corresponding color (colors that are perceived identical under different light source), many authors have tested how compute adaptation coefficient and in which space von Kries model work best [50, 38, 37].

Often, chromatic adaptation resumes to white balancing and the so called grey world assumption [76] is done (average of R, G and B color channel at edges are equal).

Tone mapping refers to the operation of modifying the value of tones in a image to make it more pleasant. This operation is particularly challenging when several dynamic range of the image is considered from the acquisition to the restitution. Generally, digital color sensor are linear with light intensity but result in a poor looking image when display or print because display of printer have intrinsic non linearity and because human processing if non linear with intensity. Tone mapping operators could be static and global, meaning that a unique non linear function applies identically for all the pixel of the image. It could also be dynamic and local depending of the local content of the scene over time [51, 59]. Dynamic local tone mapping operators could in principle include all adaptive operations such as chromatic adaptation.

There are several examples of the use of retina's models for tone mapping [58, 35, 65, 53, 59] that could serves as illustration of the modeling of retinal processing.

1.5
Conclusion

In the past, computer vision was mostly concerned with physical property of light and how it can be sensed by artificial material. Often, the final user (ex: human for photography) of image taken was not considered and we think that all the default of rendering or interpreting an image was due to a bad image acquisition system. This way of seeing was principally carried out by the argument that retina performs not so much processing information and that the brain, a too complicated structure to be studied mechanically, has non understanding algorithms inside.

This view is therefore discuss since nowadays. First the retina could show more complicated information processing that we previously thought. Second, the relationship between retinal processing and image processing suffer from several misunderstanding. Those understandings lied on the question about what is responsive for the behavior a human vision when he decides the quality or interprets an image. All the effort has been done in finding, in natural scene, the behavior of our perception without considering that biology has something to do with this behavior.

Such an approach have been very fruitful because whatever the object of the visual system is, for adapting to the physical input world optimizing transduction or for compensating inter-individual difference, the mechanisms of the visual system is always interesting to be discovered. But, neglecting the biological basis of our vision system and how it is compensated to enable us a shared vision, is certainly a major drawback of our contemporary model of perception. For computer color vision, we usually used the colorimetry reference observer $CIE1924V(\lambda)$ and $CIE1931xy$ as the complete description of the mathematical space where we want optimize some computation for computer vision (reproduction on media, detection, identification, ...).

Color mosaic sampling shared by the human retina and modern digital camera is typical examples of the compensation that arise in the visual system and may shape our perception [4, 5]. In color vision this is illustrated by the phenomenological difference between discrimination and appearance. Discrimination is, roughly speaking, the ability of a system to represent an incoming information. In such it has a quantitative behavior. Appearance in the opposite is a qualitative one because it is not necessary quantifiable and because it say something directly and not through some numbers in a defined metric space.

It happens in some circumstances that discrimination contour do not corresponds to frontiers of appearance [13, 30]. From an engineer point of view this is something surprising because we usually design frontier between categories with the locus of best discrimination. If we design a system, let say a robot with vision, we measure its response against an set of physical variables, optimized the processing for discrimination and we expect it improves our decision in between multiple categories of actions.

Actually, the problem of ignoring compensations for inter-individual differences in color vision as per-request for understanding visual cognition could be considered as noise in a variational context. For all ecological model of vision, noise come from the physical world and very few consider internal noise generated by neural processing. In color vision we usually consider luminance and color to be subject to the same independent, identically distributed noise. But, the fact that human cone mosaic is so different in between observers will certainly generate a differential noise in luminance and chrominance among observers.

The problem of how individual biological differences in human allow common vision remains to be solved for better explaining how we share so much agreement in vision despite our very different biological material. The answer to this question will certainly improves our knowledge of the functioning of the visual system and its underlying model. It would certainly provides efficient way to reproduce natural vision into biologically-inspired computer vision.

Bibliography

1 David Alleysson. Spatially coherent colour image reconstruction from a trichromatic mosaic with random arrangement of chromatic samples. *Ophthalmic and Physiological Optics*, 30(5):492–502, 2010.

2 David Alleysson and Jeanny Hérault. Variability in color discrimination data explained by a generic model with nonlinear and adaptive processing. *Color Research & Application*, 26(S1):S225–S229, 2001.

3 David Alleysson and David Méary. Neurogeometry of color vision. *Journal of Physiology-Paris*, 106(5):284–296, 2012.

4 David Alleysson and Sabine Süsstrunk. Spatio-chromatic ica of a mosaiced color image. In *Independent Component Analysis and Blind Signal Separation*, pages 946–953. Springer, 2004.

5 David Alleysson and Sabine Susstrunk. Spatio-chromatic pca of a mosaiced color image. In *Conference on Colour in Graphics, Imaging, and Vision*, volume 2004, pages 311–314. Society for Imaging Science and Technology, 2004.

6 David Alleysson, Sabine Süsstrunk, and Jeanny Hérault. Color demosaicing by estimating luminance and opponent chromatic signals in the fourier domain. In *Color and Imaging Conference*, volume 2002, pages 331–336. Society for Imaging Science and Technology, 2002.

7 David Alleysson, Sabine Susstrunk, and Jeanny Hérault. Linear demosaicing inspired by the human visual system. *Image Processing, IEEE Transactions on*, 14(4):439–449, 2005.

8 Jose-Manuel Alonso and Yao Chen. Receptive field. *Scholarpedia*, 4(1):5393, 2009.

9 Joseph J Atick. Could information theory provide an ecological theory of sensory processing? *Network: Computation in neural systems*, 3(2):213–251, 1992.

10 Joseph J Atick, Zhaoping Li, and A Norman Redlich. Understanding retinal color coding from first principles. *Neural computation*, 4(4):559–572, 1992.

11 Joseph J Atick and A Norman Redlich. Towards a theory of early visual processing. *Neural Computation*, 2(3):308–320, 1990.

12 Joseph J Atick and A Norman Redlich. What does the retina know about natural scenes? *Neural computation*, 4(2):196–210, 1992.

13 Romain Bachy, Jérôme Dias, David Alleysson, and Valérie Bonnardel. Hue discrimination, unique hues and naming. *JOSA A*, 29(2):A60–A68, 2012.

14 Horace B Barlow. Possible principles underlying the transformation of sensory messages. *Sensory communication*, pages 217–234, 1961.

15 B.E. Bayer. Color imaging array, July 20 1976. US Patent 3,971,065.

16 William Beaudot, Patricia Palagi, and Jeanny Hérault. Realistic simulation tool for early visual processing including space, time and colour data. In *New Trends in Neural Computation*, pages 370–375. Springer, 1993.

17 William HA Beaudot. Adaptive spatiotemporal filtering by a neuromorphic model of the vertebrate retina. In *Image Processing, 1996. Proceedings., International Conference on*, volume 1, pages 427–430. IEEE, 1996.

18 Anthony J Bell and Terrence J Sejnowski.

The "independent components" of natural scenes are edge filters. *Vision research*, 37(23):3327–3338, 1997.

19 Ethan A Benardete and Jonathan D Victor. An extension of the m-sequence technique for the analysis of multi-input nonlinear systems. In *Advanced methods of physiological system modeling*, pages 87–110. Springer, 1994.

20 Alexandre Benoit, Alice Caplier, Barthélémy Durette, and Jeanny Hérault. Using human visual system modeling for bio-inspired low level image processing. *Computer vision and Image understanding*, 114(7):758–773, 2010.

21 SA Billings and SY Fakhouri. Identification of systems containing linear dynamic and static nonlinear elements. *Automatica*, 18(1):15–26, 1982.

22 Pierre Bouguer. *Essai d'optique, sur la gradation de la lumiere*. Claude Jombert, 1729.

23 Gershon Buchsbaum and A Gottschalk. Trichromacy, opponent colours coding and optimum colour information transmission in the retina. *Proceedings of the Royal society of London. Series B. Biological sciences*, 220(1218):89–113, 1983.

24 P Buser, M Imbert, and RH Kay. Vision. 1992.

25 EJ Chichilnisky and Brian A Wandell. Trichromatic opponent color classification. *Vision research*, 39(20):3444–3458, 1999.

26 Giovanna Citti and Alessandro Sarti. Neuromathematics of vision. *AMC*, 10:12.

27 Laurent Condat. A generic variational approach for demosaicking from an arbitrary color filter array. In *Image Processing (ICIP), 2009 16th IEEE International Conference on*, pages 1625–1628. IEEE, 2009.

28 Philippe Crama and Johan Schoukens. Hammerstein–wiener system estimator initialization. *Automatica*, 40(9):1543–1550, 2004.

29 Dennis M Dacey, Beth B Peterson, Farrel R Robinson, and Paul D Gamlin. Fireworks in the primate retina: in vitro photodynamics reveals diverse lgn-projecting ganglion cell types. *Neuron*, 37(1):15–27, 2003.

30 MV Danilova and JD Mollon. Foveal color perception: minimal thresholds at a boundary between perceptual categories.

Vision research, 62:162–172, 2012.

31 Dawei W Dong and Joseph J Atick. Statistics of natural time-varying images. *Network: Computation in Neural Systems*, 6(3):345–358, 1995.

32 Eric Dubois. Frequency-domain methods for demosaicking of bayer-sampled color images. *IEEE Signal Processing Letters*, 12(12):847, 2005.

33 Christina Enroth-Cugell and John G Robson. The contrast sensitivity of retinal ganglion cells of the cat. *The Journal of physiology*, 187(3):517–552, 1966.

34 Gustav Theodor Fechner. *Elemente der psychophysik*, volume 2. Breitkopf & Härtel, 1907.

35 James A Ferwerda, Sumanta N Pattanaik, Peter Shirley, and Donald P Greenberg. A model of visual adaptation for realistic image synthesis. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pages 249–258. ACM, 1996.

36 David J Field. Relations between the statistics of natural images and the response properties of cortical cells. *JOSA A*, 4(12):2379–2394, 1987.

37 Graham D Finlayson and Sabine Süsstrunk. Performance of a chromatic adaptation transform based on spectral sharpening. In *Color and Imaging Conference*, volume 2000, pages 49–55. Society for Imaging Science and Technology, 2000.

38 Graham D Finlayson, Javier Vazquez-Corral, Sabine Süsstrunk, and Maria Vanrell. Spectral sharpening by spherical sampling. *JOSA A*, 29(7):1199–1210, 2012.

39 Tim Gollisch and Markus Meister. Eye smarter than scientists believed: neural computations in circuits of the retina. *Neuron*, 65(2):150–164, 2010.

40 Bahadir K Gunturk, Yucel Altunbasak, and Russell M Mersereau. Color plane interpolation using alternating projections. *Image Processing, IEEE Transactions on*, 11(9):997–1013, 2002.

41 H Keffer Hartline. The response of single optic nerve fibers of the vertebrate eye to illumination of the retina. *Am J Physiol*, 121(2):400–415, 1938.

42 David J Heeger, Eero P Simoncelli, and J Anthony Movshon. Computational

models of visual processing. In *Proceedings of the National Academy of Science*, volume 23, pages 623–627, 1996. Link.

43 J. Hérault. *Vision: Images, Signals and Neural Networks - Models of Neural Processing in Visual Perception*. World Scientific.

44 Jeanny Hérault and Barthélémy Durette. Modeling visual perception for image processing. In *Computational and Ambient Intelligence*, pages 662–675. Springer, 2007.

45 Keigo Hirakawa and Thomas W Parks. Adaptive homogeneity-directed demosaicing algorithm. *Image Processing, IEEE Transactions on*, 14(3):360–369, 2005.

46 Alan L Hodgkin and Andrew F Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of physiology*, 117(4):500, 1952. Link.

47 Ron Kimmel. Demosaicing: image reconstruction from color ccd samples. *Image Processing, IEEE Transactions on*, 8(9):1221–1228, 1999.

48 JJ Koenderink, WA Van de Grind, and MA Bouman. Opponent color coding: A mechanistic model and a new metric for color space. *Kybernetik*, 10(2):78–98, 1972.

49 Toko Kohei, Jinhui Chao, and Reiner Lenz. On curvature of color spaces and its implications. In *Conference on Colour in Graphics, Imaging, and Vision*, volume 2010, pages 393–398. Society for Imaging Science and Technology, 2010.

50 KM Lam. *Metamerism and colour constancy*. PhD thesis, University of Bradford, 1985.

51 Edwin H Land. Recent advances in retinex theory and some implications for cortical computations: color vision and the natural image. *Proceedings of the National Academy of Sciences of the United States of America*, 80(16):5163, 1983.

52 Michael S Landy and J Anthony Movshon. *Computational Models of Visual Processsing*. The MIT Press, 1991. Link.

53 Gregory Ward Larson, Holly Rushmeier, and Christine Piatko. A visibility matching tone reproduction operator for high dynamic range scenes. *Visualization and Computer Graphics, IEEE Transactions on*, 3(4):291–306, 1997.

54 Te-Won Lee, Thomas Wachtler, and Terrence J Sejnowski. Color opponency is an efficient representation of spectral properties in natural scenes. *Vision research*, 42(17):2095–2103, 2002.

55 Xin Li, Bahadir Gunturk, and Lei Zhang. Image demosaicing: A systematic survey. In *Electronic Imaging 2008*, pages 68221J–68221J. International Society for Optics and Photonics, 2008.

56 Yue M Lu, Mina Karzand, and Martin Vetterli. Demosaicking by alternating projections: theory and fast one-step implementation. *Image Processing, IEEE Transactions on*, 19(8):2085–2098, 2010.

57 Rastislav Lukac. *Single-sensor imaging: methods and applications for digital cameras*. CRC Press, 2012.

58 Rafal Mantiuk, Karol Myszkowski, and Hans-Peter Seidel. A perceptual framework for contrast processing of high dynamic range images. *ACM Transactions on Applied Perception (TAP)*, 3(3):286–308, 2006.

59 Laurence Meylan, David Alleysson, and Sabine Süsstrunk. Model of retinal local adaptation for the tone mapping of color filter array images. *JOSA A*, 24(9):2807–2816, 2007.

60 Bruno A Olshausen and David J Field. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *NATURE*, 381:607, 1996.

61 Bruno A Olshausen and David J Field. Sparse coding of sensory inputs. *Current opinion in neurobiology*, 14(4):481–487, 2004.

62 Gustav Osterberg. *Topography of the layer of rods and cones in the human retina*. Nyt Nordisk Forlag, 1935.

63 Jean Petitot. The neurogeometry of pinwheels as a sub-riemannian contact structure. *Journal of Physiology-Paris*, 97(2):265–309, 2003.

64 Rajeev Ramanath, Wesley E Snyder, Youngjun Yoo, and Mark S Drew. Color image processing pipeline. *Signal Processing Magazine, IEEE*, 22(1):34–43, 2005.

65 Erik Reinhard and Kate Devlin. Dynamic range reduction inspired by photoreceptor

physiology. *Visualization and Computer Graphics, IEEE Transactions on*, 11(1):13–24, 2005.

66 Alejandro Ribes and Francis Schmitt. Linear inverse problems in imaging. *Signal Processing Magazine, IEEE*, 25(4):84–99, 2008.

67 Bernhard Riemann. Ueber die hypothesen, welche der geometrie zu grunde liegen. *Habilitationsschrift, Abhandlungen der Königlichen Gesellschaft der Wissenschaften zu Göttingen*, 13, 1854. Link. Trans. English by Clifford W. On the hypothesis which lie at the basis of Geometry (1854), Nature, 1873, Link. Trad. Française par Houel J. Sur les hypothèses qui servent de fondement à la géométrie Link.

68 R.W. Rodieck. *The First Steps in Seeing*. Sinauer Associates, Incorporated, 1998.

69 Daniel L Ruderman, Thomas W Cronin, and Chuan-Chin Chiao. Statistics of cone responses to natural images: Implications for visual coding. *JOSA A*, 15(8):2036–2045, 1998.

70 Erwin Schrödinger. Grundlinien einer theorie der farbenmetrik im tagessehen. *Annalen der Physik*, 368(22):481–520, 1920.

71 RM Shapley and JD Victor. Nonlinear spatial summation and the contrast gain control of cat retinal ganglion cells. *The Journal of physiology*, 290(2):141–161, 1979.

72 Robert Shapley and Christina Enroth-Cugell. Visual adaptation and retinal gain controls. *Progress in retinal research*, 3:263–346, 1984.

73 Stanley S Stevens. On the psychophysical law. *Psychological review*, 64(3):153, 1957.

74 WS Stiles. A modified helmholtz line-element in brightness-colour space. *Proceedings of the Physical Society*, 58(1):41, 1946.

75 Dharmesh R Tailor, Leif H Finkel, and Gershon Buchsbaum. Color-opponent receptive fields derived from independent component analysis of natural images. *Vision Research*, 40(19):2671–2676, 2000.

76 Joost Van De Weijer and Theo Gevers. Color constancy based on the grey-edge hypothesis. In *Image Processing, 2005.*

ICIP 2005. IEEE International Conference on, volume 2, pages II–722. IEEE, 2005.

77 Hans van Hateren. A cellular and molecular model of response kinetics and adaptation in primate cones and horizontal cells. *Journal of Vision*, 5(4), 2005. Link.

78 J Hans van Hateren and Dan L Ruderman. Independent component analysis of natural image sequences yields spatio-temporal filters similar to simple cells in primary visual cortex. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 265(1412):2315–2320, 1998.

79 JH Van Hateren. Encoding of high dynamic range video with a model of human cones. *ACM Transactions on Graphics (TOG)*, 25(4):1380–1399, 2006. Link.

80 JH Van Hateren and HP Snippe. Simulating human cones from mid-mesopic up to high-photopic luminances. *Journal of Vision*, 7(4), 2007.

81 V. Volterra. *Theory of Functionals and of Integral and Integro-differential Equations*. Dover Books on Mathematics Series. Dover Publications, 2005.

82 Hermann Von Helmholtz. *Handbuch der physiologischen Optik: mit 213 in den Text eingedruckten Holzschnitten und 11 Tafeln*, volume 9. Voss, 1866.

83 JJ Vos and PL Walraven. An analytical description of the line element in the zone-fluctuation model of colour vision—i. basic concepts. *Vision research*, 12(8):1327–1344, 1972.

84 Michael A Webster and JD Mollon. Changes in colour appearance following post-receptoral adaptation. *Nature*, 349(6306):235–238, 1991.

85 Frank S Werblin and Botond M Roska. Parallel visual processing: A tutorial of retinal function. *International Journal of Bifurcation and Chaos*, 14(02):843–852, 2004.

86 Adrien Wohrer and Pierre Kornprobst. Virtual retina: a biological retina model and simulator, with contrast gain control. *Journal of computational neuroscience*, 26(2):219–249, 2009.

87 Qasim Zaidi and Arthur G Shapiro. Adaptive orthogonalization of opponent-color signals. *Biological cybernetics*, 69(5-6):415–428, 1993.