A Dual-Channel System For Human Color Vision – A Theoretical Approach

Abstract
Vision starts with hyperpolarization of photoreceptor cells after converting light into electric signals through signal transduction cascades. Those electric signals will undergo a series of processing as they pass through a chain of cells found in the retina before ending up in the visual cortex. Although research has accumulated a large amount of data about vision, our understanding of vision is still quite limited and largely fragmented. Here I present a high-level view to explain how the initial visual signals are transformed into their final formats that the central brain uses to perceive the world around us. At the same time I have proposed concrete functions for the various cell types in the retina. The rate of converting light into electric signals in a photoreceptor is controlled by the wavelength of light, and the electric signal ν is the result of the rate times the intensity of light. The wavelength and intensity are the dual properties of visible light, and each property contributes to the ν in its own right. If a pair of M and L cones are exposed to lights of the same dual properties, they form a lambda pair and generate different νs, lambda νs. The retina seems to have developed a mechanism to extract the dual properties of light from the lambda νs. The ratio of lambda νs, denoted as λ, is a constant independent of changes in intensities. Hence, λ is the neural form of wavelength in color vision. Unlike λ, the intensity signal, denoted as ι, seems to require additional processing to compensate for spectral variations attributed to wavelengths. In the fovea, a single cone cell forks its ν into two bipolar cells. One bipolar cell keeps the original negative sign of the ν, and the other reverses the negative sign to positive by different glutamate receptors. The purpose of this signal divergence at the cone axon terminals is to separate the dual properties of light into the dual channel system, one channel for processing λ, and the other for processing ι. Horizontal cells sit underneath photoreceptors and form a sensor network. Only positive bipolar cells make synaptic contacts with horizontal cells. Therefore, one concrete function for the horizontal cells seems to be controlling the entry of the initial visual signals into the positive bipolar cells. In contrast, νs that flow into the negative bipolar cells are unaltered without interference from horizontal cells, as any change would change the intrinsic neural wavelength λ calculated from the lambda νs. It is proposed that the positive bipolar cells form the positive channel to process and deliver ι, and the negative bipolar cells form the negative channel to process and deliver λ. Such a functional and structural separation of the visual pathway into two channels ensures that one signal will be processed independent of the other with the greatest flexibility and accuracy. A working model was proposed to explain how horizontal cells perform lambda pairing through executing the "sandwich algorithm" via a special cellular structures ribbon triads. The result was the division of the retinal cone mosaic into different clusters, each of which is exposed to light of the same dual properties. The final visual signals λ and ι are delivered via the dual channels to the visual cortex where they are recombined to make perception of millions of colors a reality. The visual signal processing and its underlying cellular structures can be described and predicted with a set of four equations. Key words: color vision, depolarization, hyperpolarization, phototransduction, wavelength, intensity, photoreceptor, rod, S-cone, M-cone, L-cone, retina, fovea, bipolar cell, ganglion cell, horizontal cell, amacrine cell, neurotransmitter, glutamate, retinal, H-cluster, sandwich algorithm.

1. Introduction

The human eye is the finest camera invented and has been whetted relentlessly by mother nature over millions of years of evolution. It captures natural themes with such a high degree of fidelity and allows us to enjoy the world in full color. Over the years, great effort has been devoted to understand the cellular and neurological bases of vision, especially the retina of eye. For the simplicity of description, the structure of the retina is roughly divided into three layers (Figure 1). The inner layer, facing the pupil of eye, consists of ganglion cells, whose axons form optical nerves that carry highly processed visual information to the central brain. The dendrites of ganglion cells receive signal inputs from the middle layer consisting of bipolar cells, which receive visual information from the outer photoreceptor layer. It's the photoreceptor layer that generates the raw electric potentials upon exposure to light.
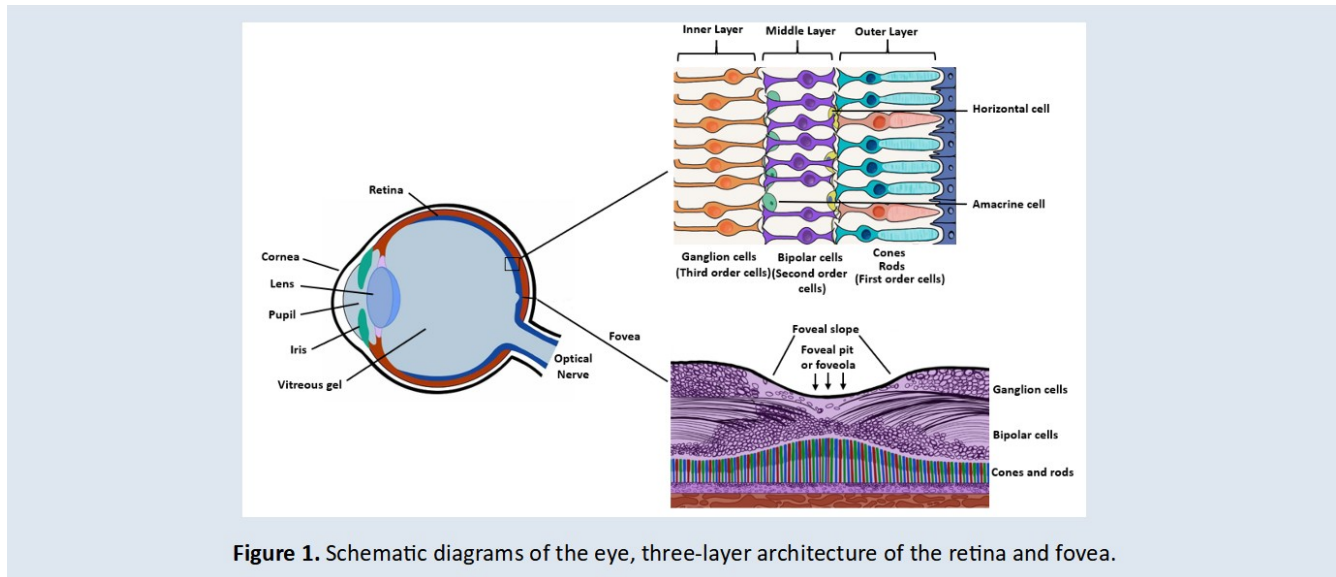

Figure 1. Schematic diagrams of the eye, three-layer architecture of the retina and fovea.

The tissue space between the photoreceptor layer and bipolar cell layer is called the outer plexiform layer(OPL), and the tissue space between the middle and inner layers where bipolar cell axons and ganglion cell dendritic terminals connect is called the inner plexiform layer(IPL). IPL is divided into two functionally discrete sub-laminae. The sub-lamina a is located at the middle layer side, and the sub-lamina b located at the inner layer side. The axon terminals of a given bipolar cell type and the dendritic terminals of a given ganglion cell type usually end either in sub-lamina a or sub-lamina b (mono-stratified), although some of these cells span their axon or dendritic tree into both sub-lamina a and sub-lamina b (bistratified). Such a clear division of tissue space is of great biological significance.

The fovea is a small depression of about 1.5 mm in diameter in the central retina and is a functionally and morphologically specialized structure in primates only. In humans, the flat center of the fovea, called foveola or foveal pit, sees about 1.5 degrees of the central visual field, which is the ultimate view with the highest acuity that allows us to read and discriminate very tiny differences. Vision from the foveal area will be the focus of this paper.

Two types of photoreceptor cells are present in the photoreceptor layer, cone and rod. Their light reactivity comes from the light absorbing chromophore, 11-cis retinal-opsin complex. When 11-cis retinal couples to different types of opsin molecules, it absorbs light in different wavelength ranges. All of primate color vision comes down to three distinct types of cones, each of which contains a unique type of opsin. S (short) cone is responsible for light in the short wavelength range, M (medium) cone for light in the medium wavelength range, and L (long) cone for light in the long wavelength range.

The opsin present in rod cell is called rhodopsin. Because there is only one form of rhodopsin for all rod cells, the rod cell is not color-capable. Figure 2 shows the normalized absorption spectral curves for the three cone types plus the rod cells. There are considerable overlaps in the wavelength ranges in which the three cone types react to light and generate the initial visual signals.
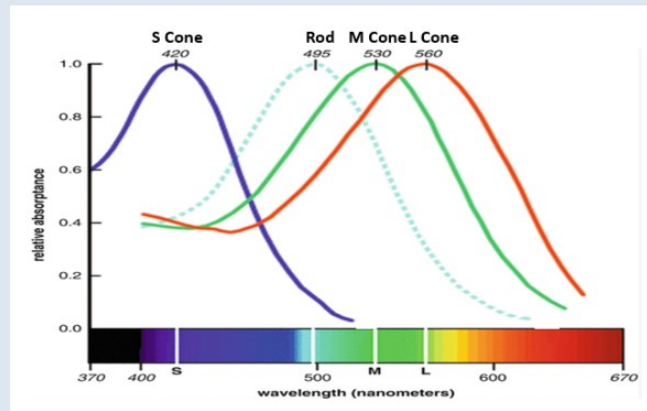

Figure 2. The normalized absorption spectra of the three cones plus rod.

Research shows that a single cone type can be excited to generate a particular electric potential by light of different combinations of wavelength and intensity. By increasing the intensity, a weakly absorbed wavelength can evoke a cone response that will be indistinguishable from the response evoked with strongly absorbed wavelength at a lower intensity, suggesting a single cone type is not sufficient to differentiate a response change in wavelength from a response change in intensity. The color visual information must be extracted from the responses across different types of cones.

In this paper, by taking a theoretical approach, I will present mechanisms to explain how primate color vision is achieved in retina. Hopefully it will allow us to think of vision from a new angle with a goal to gain a better understanding of color vision and facilitate the interpretation of research data.

2. Wavelength and Intensity Are the Dual Properties of Light That Determine Color
In photochemistry, the chance that a photoreaction will occur depends on the wavelength of the light that the reactant is sensitive to, and the degree of the photoreaction triggered by that wavelength is determined by the intensity of that light. Therefore, the wavelength and intensity are the dual properties of light that determine if a photoreaction will occur and occur in what degree, which, in color vision, determine the color of light in the central brain. Figure 3 shows that the color green changes its tone as the light intensity decreases, while the green wavelength is unchanged. Clearly wavelength and intensity are two equally important physical properties of the color.
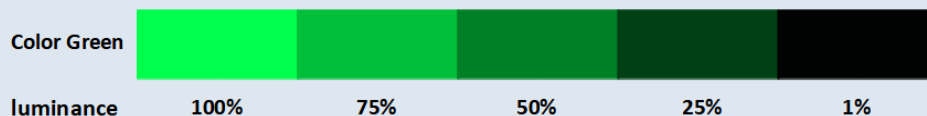

Figure 3. The rectangles represent the tones that change from light green to dark green as their luminance decreases from 100% to 1% while green wavelength is unchanged.

Vision is essentially a photoreaction followed by the phototransduction pathway to convert light energy into electric potential Vs across the membrane of a cone cell, referred to as hyperpolarization – the very

initial vision signals. In the dark, a photoreceptor is partially depolarized and releases neurotransmitter glutamate to impact the downstream neurons. By contrast, a photoreceptor reduces its glutamate release upon hyperpolarization. This seemingly simple and trivial consequence of hyperpolarization as a result of phototransduction encapsulates all the information needed for color vision.

The electric potential $\mathcal{V}$ is the only product that a cone produces in response to light exposure, and its size is determined by the absorption spectral curves of the cone type as well as the number of photons that trigger the photoreactions. Red light is poorly absorbed by all cone types, but to our eye, it is as brilliant as strongly absorbed colors green and yellow. The retina must not only capture the light in the form of $\mathcal{V}$s, but also possess mechanisms to convert the $\mathcal{V}$s into normalized signals that truly reflect the wavelength and intensity of the incoming light in an unbiased way. How does the retina extract visual information from light responses generated by different cone types? How does the retina distinguish if a $\mathcal{V}$ is the result of stimulation of green or red light or light of other mixed wavelengths? How does the retina normalize signals produced from photoreactions involving all the wavelengths across the visible spectrum to defy great variations in spectral absorption? These are very fundamental questions that must be accounted for if we want to understand color vision.

The photons become irrelevant after photoreactions, and the rate of the phototransduction pathway will depend only on the number of photons that trigger the photoreactions, which further determines the level of cone output measured in the amount of reduction in glutamate release. If the efficiency of photoreactions or phototransduction pathway fluctuated, the sizes of electric potentials $\mathcal{V}$s at the cone bases would fluctuate as well, resulting in fluctuation in the amount of reduction in glutamate released into the downstream neurons. When this happened, it would be impossible for the central brain to interpret the end results of phototransduction in an accurate way. Therefore, all the reactions in the visual signal production and processing must be consistent and steady. Borrowed from mathematics, the consistent efficiency of phototransduction pathway is expressed as a coefficient c.

The rate $\mathbf{r}$ at which the photons are absorbed by 11-cis retinal-opsin complex is determined by the wavelength of the photons. The $\mathbf{r}$ will be smaller if the photon carries a wavelength farther away from the absorption peak, but it will be a constant for a particular wavelength in a particular cone type. The electric potential $\mathcal{V}$ or hyperpolarization could be approximately described as the product of the rate $\mathbf{r}$ times $\mathbf{I}$, the number of photons or the intensity of light in a unit of time, times c, the coefficient of phototransduction pathway shown below:

$$\mathcal{V} = \mathbf{r}c\mathbf{I}$$

Since $\mathbf{r}$ and $\mathbf{c}$ are constants, they are combined into another constant $\mathbf{R}$ to have the following equation:

$$\mathcal{V} = \mathbf{R}\mathbf{I} \tag{1}$$

Wavelength appears in the constant $\mathbf{R}$ only implicitly, making it an intrinsic property of light. In contrast, intensity determines the value $\mathcal{V}$ of a given wavelength in a positive fashion by directly being part of the calculation. Above equation also shows that the same level of hyperpolarization can be achieved with different combinations of wavelength and intensity. A single cone type is unable to differentiate a change due to wavelength or due to intensity.

A pair of M-cone and L-cone are named lambda pair when they are exposed to the same light, and the electric potentials $\mathcal{V}$s produced by the lambda pair are called lambda $\mathcal{V}$s. By applying Equation (1) to the lambda pair, we have the following equation pair:

$$\mathcal{V}_M = R_M I$$
$$\mathcal{V}_L = R_L I$$

By taking the division of two equations, we have the equation below:

$$\frac{\mathcal{V}_M}{\mathcal{V}_L} = \frac{R_M I}{R_L I} = \frac{R_M}{R_L}$$

The ratio of the lambda $\mathcal{V}$s is equal to the ratio of R_M to R_L, the result of which is a constant. Therefore, the ratio of the lambda $\mathcal{V}$s is a constant as well. Since $\mathcal{V}$s are the end products of phototransduction and the very initial signals of color vision, the constant nature of the ratio suggests that it's the neural form of wavelength that the visual system generates to represent the physical wavelength of the incoming light. An inference is that the central brain uses this neural form of wavelength as one input to perceive colors. The neural form of wavelength is denoted λ. We have a definition of λ:

$$\lambda = \frac{R_1}{R_2} = \frac{\mathcal{V}_1}{\mathcal{V}_2} \tag{2}$$

Subscripts 1 and 2 in Equation (2) mean M-cone or L cone in a lambda pair. The constant λ is light intensity independent, allowing the visual system to achieve color constancy, the ability to perceive the same color as the same color all the time. Figure 4 shows two highly hypothesized response lines, blue for M-cone and orange for L-cone. The yellow line that represented the ratio of L-cone to M-cone responses was flat and constant when both cone responses were linear. It displayed a unique intensity-independent nature like physical wavelength that is independent of the intensity of the light.
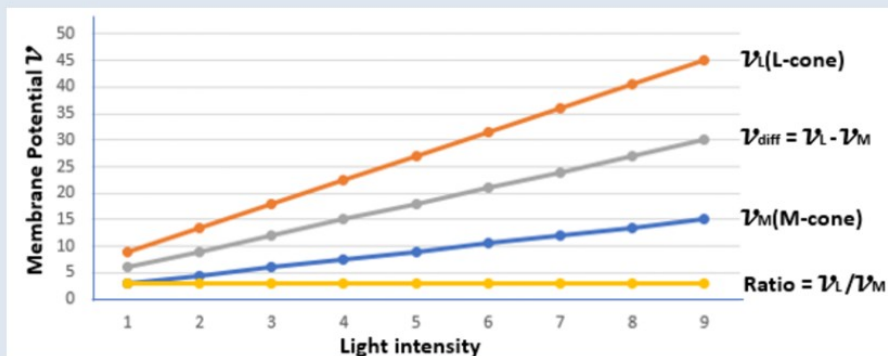
Figure 4. Hypothesized light responses by an M-cone and an L-cone in terms of electric voltage or membrane potentials. V-Diff (gray) and Ratio (yellow) are derived from the original data lines.

If the retina keeps the order of division consistent when deriving the λ, for example, $\mathcal{V}_M$ over $\mathcal{V}_L$, the λ values can represent almost all wavelengths that the M-cone and L-cone respond to. If $\mathcal{V}_M > \mathcal{V}_L$, all λs

are for blue side wavelengths; if $\mathcal{V}_L > \mathcal{V}_M$, all λs are for red side wavelengths. They are separated by the λ derived from the condition $\mathcal{V}_L = \mathcal{V}_M$.

Based on equation (1), the intensity **I** could be calculated only if $\mathcal{V}$ and **R** were known. However, **R** is conceptual and doesn't really mean anything to the retina. It isn't possible for the retina to accurately determine the intensity from the measurable end products $\mathcal{V}$s. Nevertheless, the visual system must convert the intensity of light into a neural form as it does for the wavelength. The neural form of intensity is denoted **i**, and it must be the highly accurate representation of the physical intensity of the incoming light, which the central brain must need to achieve a full perception of colors. Neural λ and neural **i** are the dual signals for color vision, corresponding to the dual properties of light – wavelength and intensity.

How does the retina extract intensity information from the end products $\mathcal{V}$s? The neural λ is an absolute value derived from the lambda $\mathcal{V}$s, but the neural **i** could be an approximation only. If we take the differences $\mathcal{V}_{diff}$ of the lambda $\mathcal{V}$s, and plot them against the increasing amount of intensity, we got the gray line shown in Figure 4. It's clear that the $\mathcal{V}_{diff}$ of a lambda pair is linear as well against the increasing amount of intensity. By applying Equation (1) to M-cone and L-cone, we have the following equation after doing the $\mathcal{V}$ differences and rearrangement:

$$ I = \frac{\mathcal{V}_1 - \mathcal{V}_2}{R_1 - R_2} = \frac{\mathcal{V}_1}{\Delta R} - \frac{\mathcal{V}_2}{\Delta R} $$

Though ΔR is unknown to the retina, it is constant for incoming light of fixed wavelength. Appearance of ΔR implies that both $\mathcal{V}_1$ and $\mathcal{V}_2$ can't be used directly out of cones for neural **i**, but must undergo a certain change to take ΔR into consideration. ΔR is called lambda effect. The lambda effect is purely due to spectral variations attributed to wavelengths. If the retina resolved the lambda effect on a $\mathcal{V}$, it gets a new $\mathcal{V}$, denoted $\mathcal{V}_\lambda$:

$$ \mathcal{V}_\lambda = \frac{\mathcal{V}}{\Delta R} \qquad (3) $$

We use $\mathcal{V}_{\lambda 1}$ and $\mathcal{V}_{\lambda 2}$ to express the lambda $\mathcal{V}$s after resolving the lambda effect ΔR. By ignoring the lambda effect, the $\mathcal{V}_{diff}$ of a lambda pair is a reasonable estimate of neural **i**. By removing the large absorption variations, $\mathcal{V}_{diff}$ as neural **i** reflects more closely the true intensity of the incoming light across the entire visible spectrum. With lambda $\mathcal{V}_\lambda$s, neural **i** could be calculated to the highest possible accuracy shown in Equation (4):

$$ i = \mathcal{V}_{\lambda 1} - \mathcal{V}_{\lambda 2} \qquad (4) $$

The Equations (3) and (4) indicate that the retina must do extra work to harness transmembrane potentials $\mathcal{V}$s before it could use them for neural **i**. In contrast, generation of λ is a straight forward operation using $\mathcal{V}$s directly from cones.

Equations (2) and (4) have effectively eliminated the impact of great spectral variations on λ and **i**, and as a result, strongly absorbed light didn't have any advantage over weakly absorbed light after they are

transformed into ultimate neural visual signals. The implications of lambda effect **ΔR** and Equations (2) and (4) are that the retina must contain a super cellular structure to resolve **ΔR** and two separate signal channels to process and convey **λ** and **ɩ.**

By separating visual signals into **λ** and **ɩ,** the retina generates and delivers far more detailed visual information to the central brain, where colors that differ only subtly in wavelength and/or intensity can be distinguished to form the real mirror image of the physical world. In this way, the vision system has achieved an extraordinary capability to perceive millions of colors through combinatorial power of **λ** and **ɩ. λ** is of chromatic nature, while the intensity of light or **ɩ** is of achromatic nature. An **ɩ** can act on any **λ** to decrease or increase a color tone specific to that **λ** as shown in Figure 3. If it is the case, the chromatic **λ**s allow the visual cortex to render base colors, while achromatic **ɩ**s allow the visual cortex to modulate the tones of colors set by the **λ**s.

Hyperpolarization of photoreceptors as the result of photon stimulation may be of great biological significance. Visual system must be able to be responsive to and differentiate light stimulation ranging from very weak to strong. Action potentials as the result of depolarization are largely responses that are threshold controlled and therefore unable to illicit responses from sub-threshold stimulation from the low light environment. Single wavelength light is limited roughly to between 400 nm and 700 nm, but multi-wavelength light is almost unlimited, especially when the intensity of each wavelength differs. Action potentials as the result of depolarization are too low resolution for the eye to distinguish stimuli from a million of colors, which can be different only minusculely in wavelength and intensity. The cone system requires mechanisms to generate **𝒗**s that are granular and broad enough to discriminate lights that differ in all degrees. It seems true that threshold independent hyperpolarization has the advantage to generate linear changes in **𝒗** to precisely reflect the wavelength composition and the overall intensity of the incoming light.

3. Visual Pathway – a Dual-Channel System for the Dual Properties of Light.
The term "algorithm" is used in computer science to describe an ordered set of operations to transform an input into an output. If an algorithm, for example, quick sort algorithm, applies to a set of data hundreds of time, it always generates the same ordered list of data using the same amount of time. In the biological system, nervous activities seem to be controlled by certain algorithms much like a computer program. Retina of the eye can actually be regarded as a biological computer, which converts physical light into electrical impulses and further transform the initial electrical impulses into final formats that the central brain can use to perceive light for imaging. Since the entire process of visual signal transformation is so reliable, predictable, and precise, it can be achieved only by following an pre-defined ordered set of operations in a strict manner. We termed such a set of operations neural algorithm. Neural algorithm, just like a computer algorithm, when applied to a specific neural task, it always generates the same result from the same set of input using the same amount of time. Without neural algorithms, we would see the same color differently every time we perceive that color. Neural algorithms are important to our nervous system just like computer algorithms are important to computer programs.

The retina consists of three layers of cells and between each layer there are also accessory cell types. Horizontal cells serve the visual pathway at OPL, and amacrine cells play roles at IPL. All cell types after photoreceptors display complicated morphology and exist in many subtypes, and each subtype can be characterized with its own unique complexity in morphology and cellular localization. They express different neurotransmitter receptors and release different neurotransmitters to act on the recipient cells with distinct effects. More importantly, they are entangled to form messy intercellular networks. Their true faces and functional details are still waiting to be revealed.

3.1 Bipolar Cells Are the Gates to the λ and ɩ Channels
The bipolar cells exist between photoreceptor layer and ganglion cell layer and make connections to cells in the both layers via synapses and gap junctions. They relay visual signals to the ganglion cells from the photoreceptors. In the fovea, one bipolar cell receives input from a single photoreceptor, but not vice versa. One cone cell transmits output to two bipolar cells. These two bipolar cells belong to two types. The first type is called invaginating bipolar cell or IBC because a dendritic terminal of it is sandwiched between two dendritic terminals from two horizontal cells, forming a special structure called an invagination at the bottom of the cone axon terminal (see Figure 5). The second type is called flat bipolar cell or FBC because its dendritic terminals make flat synaptic contacts at area next to the invagination. IBC and FBC respond to light stimulation in completely opposite manners.

There are two main types of glutamate receptors expressed on the dendritic tips of bipolar cells: flat bipolar cells express ionotropic AMPA/kainate glutamate receptors or iakGluRs, while invaginating bipolar cells express metabotropic glutamate receptors or mGluRs. The iakGluRs themselves are cation channels opened by glutamate. A decline in glutamate concentration due to reduced release closes iakGluRs, making flat bipolar cells more negative inside to become hyperpolarized. The mGluRs close a cation channel called TRPM1 upon binding of glutamate. Lower concentrations of glutamate open TRPM1, making the invaginating bipolar cells more positive inside to become depolarized. For their unique responses to neurotransmitter glutamate, flat bipolar cells are called negative or N bipolar cells, while invaginating bipolar cells are called positive or P bipolar cells.

The visual signals fork into P and N bipolar cells upon leaving the photoreceptors both in fovea and peripheral. However, there is one difference: each bipolar cell in the peripheral gets its input from multiple cones instead of a single cone as in the fovea. The axon terminals of a P bipolar cell end up in sub-lamina b, where via synapses they connect with the dendrites of ganglion cells that also end up in sub-lamina b. These ganglion cells depolarize upon receiving neural signals from P bipolar cells; hence they are positive or P ganglion cells. Similarly, the axon terminals of an N bipolar cell end up in sub-lamina a, where they synapse with the dendrites of ganglion cells that end up also in sub-lamina a. These ganglion cells hyperpolarize upon receiving neural signals from N bipolar cells; hence they are negative or N ganglion cells. Such distinct cellular structures and correspondingly distinct neurological properties strongly suggest that there exist two channels in the visual pathway; the positive or P channel conveys neural messages via action potential, while the negative or N channel conveys neural messages via hyperpolarization. The P channel and N channel make up the dual channels of the visual pathway for color vision.
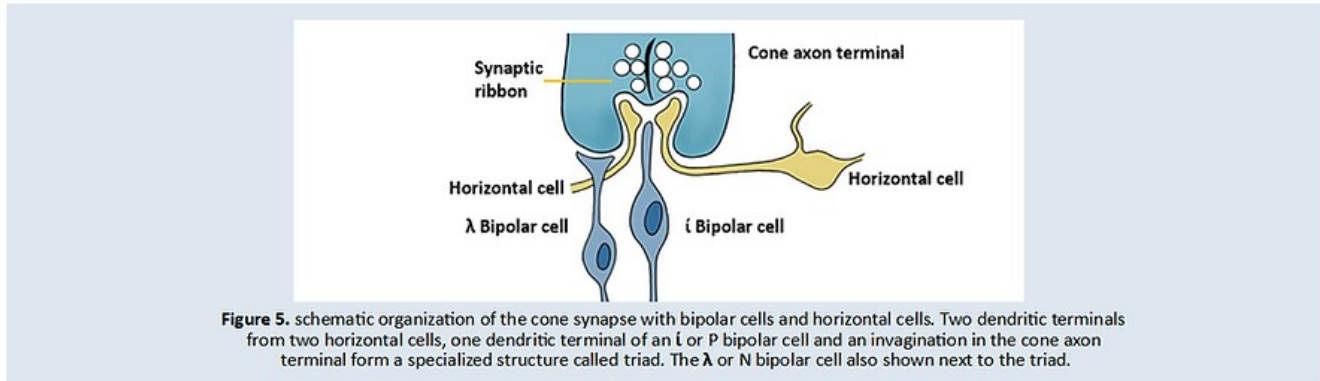
Obviously the dual-channel system is a simple mechanism that the evolution has chosen to separate a complex problem into two parts right after the raw visual signals are generated. The clear advantage to have such a dual-channel system is that each part can be processed and conducted in its own space without interfering with each other, which is what is precised predicted by the Equations (2) and (4). Evidently the dual channels are essentially the λ and ɩ channels, and the dendritic terminals of bipolar cells are the entry points where signals separate and flow into these two channels. A dual channel system provides cellular structures to keep λ signal constant as indicated by Equation (2) and allow ɩ signals to undergo extra processing to resolve the lambda effect as shown by Equations (3) and (4).

3.2 Current Understanding of Horizontal Cells in the Visual Signal Transformation Pathway
Horizontal cells are a type of interneurons, and their flat arborized dendritic terminals lie underneath the axon terminals of cones. A horizontal cell makes synaptic contacts with many photoreceptors, and a single cone cell can be contacted by several horizontal cells. Horizontal cells thus form a complex

interneuronal network with cones. In the meanwhile, they form synaptic contacts with P bipolar cells. Because of expressing iakGluRs, they hyperpolarize upon cone exposure to light.

A cone cell has multiple invaginations at its axon terminals. Each invagination houses two dendritic tips from two horizontal cells plus a dendritic tip of P bipolar cell in between. Inside the cone axon terminal is another specialized structure called ribbon. The ribbon synapses can release glutamate into the synapse cleft in a fast and steady fashion. The complex that consists of a ribbon synapse, two dendritic tips from two horizontal cells and a P bipolar cell dendritic tip is called ribbon triad (Figure 5). The ribbon triad provides a micro-environment in which the retina performs important functions during visual signal transformation.

Figure 5. schematic organization of the cone synapse with bipolar cells and horizontal cells. Two dendritic terminals from two horizontal cells, one dendritic terminal of an ℓ or P bipolar cell and an invagination in the cone axon terminal form a specialized structure called triad. The λ or N bipolar cell also shown next to the triad.

In research literature, horizontal cells play a number of roles in visual signal processing: contributing to contrast enhancement and color opponency, generating center-surround receptive fields, and providing lateral inhibition, feedback and feed-forward interactions to photoreceptors and bipolar cells. This list includes almost every major purported visual processing feature that contributes to color vision. Is it possible for a single cell type to play roles in all of them? It has been believed for over 30 years, and it continues to be believed so today. Nevertheless, cells would not assume roles imposed or given by people who have great interests in studying them.

There is a fundamental difference between neural research and general biological research. In general biological research, results obtained from individual cells or individual parts of cells are largely valid when putting back into the whole system. For example, results from in vitro transcriptional regulation study is relevant to the in vivo transcriptional regulation in question. However, results from individual neurons or a group of neurons aren't necessarily sufficient to paint a general picture of how these neurons work and function in vivo as a whole. In neural research, the sum is greater than, even much greater than, the parts. When horizontal cells are studied for their roles in visual signal transformation, it must be aware that there is no single cell type that can be a master of all, contributing to every aspect of the system. More specialized a cell type becomes, more limited functionality the cell type will have. Highly specialized horizontal cells can't be exception. Otherwise it will lead research to nowhere. Over-interpretation of research results obtained from limited conditions in a way detached from the big picture will severely obscure our understanding of the true nature of these cells.

As a result, despite decades of research, the horizontal cells' true roles in visual signal transformation are still vague and elusive. By paying attentions to their major neural characteristics while overlooking tidbit details, great attempts have been given to investigate what roles horizontal cells can play in visual signal transformation, especially their roles in the ribbon triad structure. The goal is to paint a better picture of how horizontal cells contribute to color vision in fovea.

3.3 Horizontal Cells Control the Gates to the Two Channels

The visual information flows from the cones to bipolar cells to ganglion cells. There are accessory cells to modulate the process every time the information changes hands. The direct contacts between horizontal cells and photoreceptors pose time and space constraints on the roles horizontal cells can play, which means that it's unlike for the horizontal cells to play as many roles in the vision process as they have been believed so. Furthermore, visual signals that horizontal cells receive from the cones are at the very early stage in the visual pathway, and it is impossible for the horizontal cells to figure out what role a signal from a particular cone will be in the whole visual field and take action to make necessary changes for the sake of final images, for example, contrast enhancement. Nevertheless, the fact that a horizontal cell contacts multiple cones and a cone is contacted by multiple horizontal cells reveals an important function – possible signal comparisons over the area covered by horizontal cells, which obviously requires input of information from multiple cone cells. The co-localization of cone axon terminals and horizontal cell dendrites suggests that the signals that have just reached the cone axon terminals must be processed before they enter the bipolar cells. According to equations (2), (3) and (4), generation of neural signals is a two-step task at least. The first step occurs at the OPL and the second step occurs at the IPL.

When we think of horizontal cells, we must think of their roles in how retina fulfills the first step of the task. The presence of the ribbon triad structures at the cone-P bipolar cell junctions is intriguing. Why does the retina need such a super structure? and why does the ribbon triad contain two dendritic tips from two different horizontal cells and a single dendritic tip from a P bipolar cell arranged in such a way under the ribbon? The answer obviously is to fulfill the first step of the task: matching lambda pair and resolving lambda effect. The single dendritic tip from a P bipolar cell in the ribbon triad seems to be involved only in resolving the lambda effect.

The horizontal cells are likely to be responsible for matching lambda pair by executing the match algorithm. The match algorithm instructs the matching process to gather νs from all cones covered by horizontal cells and then compare their νs to determine which pairs meet the matching requirements. It was assumed that a visual scene could be divided into clusters on the cone mosaic of the retina, and cones in each cluster were exposed to light of the same dual properties. Such a cluster would have no fixed shape and size and be heterogeneous in cone types, as it reflected a homogeneous part in the visual scene at a given moment. This kind of cone clusters is named H clusters (connected by Horizontal cells, Heterogeneous in cone types, and Homogeneous in light exposure).

The sandwich algorithm is an easy-to-conceive match algorithm. Lambda pairing requires minimum of three cones in straight line, in which the middle cone is a different type (Figure 6A). In human, cones in the foveal center are arranged in a hexagonal configuration to pack the maximum number of cones in a given area. The hexagonal configuration seems to make the sandwich algorithm more likely to work (Figure 6B). A primate photoreceptor can have many ribbon structures, and a single photoreceptor cell can be covered by multiple horizontal cells, making it possible for the sandwich algorithm to identify all lambda pairs in the fovea. The match algorithm relies on νs from all cones connected by horizontal cells, resulting in the grouping of cone cells into lambda pairs.

After cones were matched into lambda pairs, adjacent pairs could fuse into a H cluster if they shared the same νs. This process could take place in the entire fovea to delineate the cone mosaic into a number of H clusters. Any cone of one type in a H cluster could pair with any cone of different type in the same H cluster as lambda pairs. By forming a horizontal cell based sensor network to execute the sandwich algorithm, which essentially was a voltage mediated cell sorting process, every cone cell

could be mapped into a particular H cluster. When νs entered bipolar cells, cones based H clusters became either P bipolar cells based H clusters or N bipolar cells based H clusters.
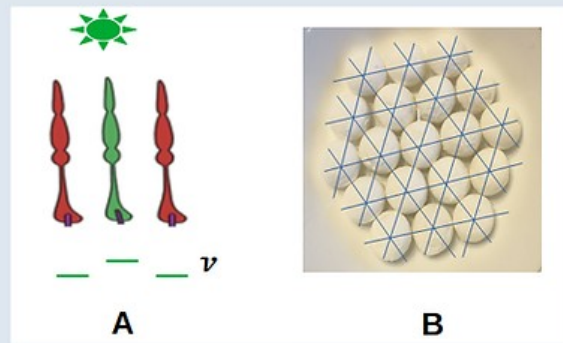

Figure 6. A). Illustration of the sandwich algorithm for matching delta pairs. If the middle M-cone is flanked by two L-cones with the identical potential νs, then M-cone must be exposed to the same light source as the two flanking L-cones. **B).** Each rice ball represents one cone, and 19 rice balls are arranged in a hexagonal configuration. In such an arrangement, each rice ball can be assessed by sandwich algorithm from all six angles.

Let's look at the signal passage from cones to N bipolar cells. N bipolar cells express iakGluRs, suggesting that they hyperpolarize upon receiving signals from the cones. Therefore, N bipolar cells essentially receive a copy of the original signal from the cone axon terminals. This preserves the sign and quantity of the signals of cone cells in the N bipolar cells, fulfilling the requirement of Equation (2) for straightforward calculation of λ. The use of unaltered lambda νs for λ indicates that there is no need for a mechanism to modulate this process. Indeed, there is no cellular structure found to carry out an additional modulation mechanism in this signal transfer process.

If the cone signals were passed to P bipolar cells like the cone signals were passed to N bipolar cells, P bipolar cells would receive a copy of cone signal except the signal sign was converted to positive. The P bipolar cells could still use raw signals to generate ι, but it's only an approximation of the intensity when the lambda effect is ignored, which clearly isn't good enough for the highest color resolution. A complex end can't be achieved with simple means. To generate ι signals, a sub-task to resolve the lambda effect $\Delta \mathbf{R}$ is likely to follow. To resolve the lambda effect $\Delta \mathbf{R}$ is to factor the $\Delta \mathbf{R}$ into the final ν_λs before νs enter P bipolar cells. $\Delta \mathbf{R}$ is a vague concept that is solely due to Equation (3). It's at the retina's discretion how to resolve the lambda effect as long as the resolution is kept consistent in all time and light conditions.

It appears that the ribbon triad structure has been evolved to resolve the lambda effect, and its presence in the retina has been predicted to satisfy Equations (3) and (4). A horizontal cell's dendritic tips express iakGluRs, which consume glutamate released from the ribbon synapse. Its direct consequence is that the glutamate pool in the synapse cleft available to the P bipolar cells is effectively reduced by a certain amount, which in turn amplifies the positive signals on the P bipolar cells by a corresponding amount. Having horizontal cell's dendritic tips in the ribbon triad seems to be a venue to resolve the lambda effect by modulating the amount of glutamate available to the P bipolar cells. The end result is turning lambda νs into lambda ν_λs.

There are more factors working in the ribbon triad micro-environment to change the concentration of glutamate in the synapse cleft, further influencing depolarization of the P bipolar cells. How exactly these factors influence each other seems to be controlled by a special neural algorithm – resolve

algorithm, which is encoded in the ribbon triad specifically for the resolution of lambda effect. At the end of the resolve algorithm, any given input lambda νs are transformed predictably and consistently into corresponding output lambda ν_λs, and the output lambda ν_λs are solely determined by the input lambda νs. A likely outcome of the resolve algorithm is that the lambda νs on the P bipolar cells have been modulated to become ν_λs whose value range is widened to certain extent, resulting in broader range of $\textit{i}$s as well. Enhanced $\textit{i}$s enable the retina to distinguish color tones due to subtle differences in intensity. It's essentially a mechanism to amplify signal differences for easier distinction. As long as the resolve algorithm is executed consistently across all range of νs, the resulting ν_λs will be consistent as well to keep the vision constant for all the sources of light. Resolution of the lambda effect can bring changes big or small to νs, but it has no impact on λ. This is clearly one of the reasons why evolution separates visual signals into two channels.

To this point, the retina has well prepared bipolar cells for subsequent signal transformation. This preparation must be completed at the time the signals diverge into λ and $\textit{i}$ channels for economical and space reasons. Horizontal cells control the gate to $\textit{i}$ channel to ensure that the initial visual signals have lambda effect resolved and that each cone and P bipolar cell are mapped into their own H clusters, making P bipolar cells ready for the second step to calculate $\textit{i}$ values. Figure 7 summarizes how the retina converts νs from a lambda pair into two separate, independent parameters λ and $\textit{i}$, and conveys them in their own channel to central brain where they are merged and translated into color.
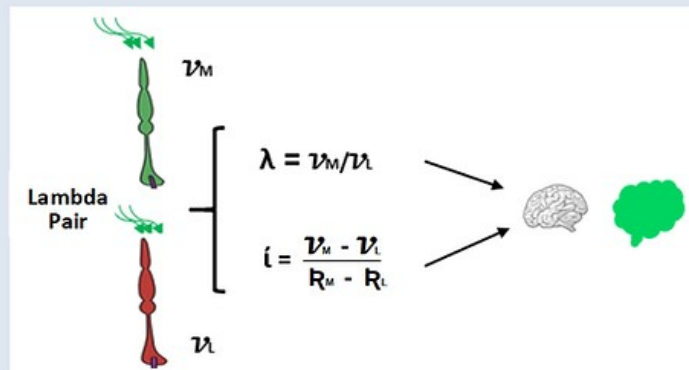

Figure 7. Summary of conversion of cross membrane electric potential νs from the two cone types (lambda pair) into neural wavelength λ and neural intensity $\textit{i}$, which are the ultimate neural signals transmitted to the brain where they are decoded into color that we perceive.

It would be interesting to see if different individuals resolve lambda effect equally well. If not, then a better ability to resolve the lambda effect could confer individuals to enjoy more vivid colors. The good thing is that if individuals never sees vivid colors due to their not-so-good ability to resolve the lambda effect, they wouldn't even know that colors could be viewed more vividly. Therefore, the red in one's eye may not be exactly identical to the red in other's eyes. As long as colors remain the same to an individual, it's alright for the colors to be viewed differently to different people.

Vision acuity varies greatly among individuals. It can be said that there seem a few major factors that could affect the vision acuity. Smaller diameters of cone cells would allow more cone cells to be packed into the hexagonal configuration in the fovea, enhancing space acuity. A balanced ratio and precise distribution of M-cones and L-cones in the fovea could increase color acuity by increasing the efficiency and accuracy of the execution of match algorithm. Lastly, a superior ability to resolve the lambda effect could greatly enhance color vision acuity.

3.4 Amacrine Cells Do the Rest of Work

Amacrine cells, like horizontal cells, are a type of interneurons in the retina. They forms a large family of many subtypes with various shapes and sizes. Their most prominent feature is the great arborization of dendritic trees to cover small to large visual fields, depending on their tissue locations. Amacrine cells are strictly either mono-stratified or bi-stratified in the IPL region, where the axon terminals of bipolar cells and the dendrites of ganglion cells interconnect. Meanwhile, they make extensive and intricate connectivity to other cells in their reaches. Amacrine cells perform important functions both in local and global scopes.

After visual signals enter bipolar cell layer, N bipolar cells with hyperpolarization signals are ready for conversion into the neural wavelength λ, while P bipolar cells with depolarization signals are ready for conversion into the neural intensity ι. Both conversions will take place independently in the sub-lamina space in their own channel with the help of amacrine cells sitting in between.

Neural λ is produced by dividing lambda νs (Equation (2)), and neural ι is produced by subtracting lambda ν_λs (Equation (4)). Here division and subtraction are operations borrowed from mathematics to make a point that νs of a lambda pair will be combined in certain ways to generate new membrane potentials in the ganglion cells. Output of λ and ι must be predictable and precise in quantity, time, and space for every input of ν and ν_λ. This two processes, similar to lambda pair matching and lambda effect resolution, must be instructed by their own algorithms – λ algorithm for λ generation, and ι algorithm for ι generation.

In fovea execution of λ algorithm occurs in sub-lamina a, where the axons of N bipolar cells contact dendrites of N ganglion cells. Execution of ι algorithm occurs in sub-lamina b, where the axons of P bipolar cells contact dendrites of P ganglion cells. λ algorithm is independent of ι algorithm, but ι algorithm might require input of signals produced by the λ algorithm. Since hyperpolarization signals reach the sub-lamina a earlier, the execution of λ algorithm could occur ahead of ι algorithm.

Contacts between bipolar cell axons and ganglion cell dendrites form a large number of synapses, as both axons and dendrites are quite arborized. When the neural signals reached the branching points on bipolar cell axons, they could be partitioned into each axon terminal. Each axon terminal could carry its own threshold to control if the signal it received sufficed to release neurotransmitters into the postsynaptic cells. Amacrine cells, integral part of the algorithms, would influence greatly the outcome of the algorithms. The central role of amacrine cells could be to either strengthen or weaken the signal that flows from bipolar cells to ganglion cells. The nature of the outcome could be further influenced by information amacrine cells received from other amacrine cells and/or bipolar cells or ganglion cells. Only survived signals could release neurotransmitters to act on postsynaptic dendrites of the ganglion cells. With their large number of subtypes to cover various visual fields in different locations, amacrine cells could exert broader influences in the visual pathway beyond λ algorithm and ι algorithm, for example, coordinating, integrating, and synchronizing visual signals from large visual areas to enhance the overall quality of color vision. The ganglion cells finally summate the signals from their dendritic trees into final bits of information in axons and transmit them to the central brain.

Although the neural division and neural subtraction could be complicated, both would follow the steps set by λ algorithm and ι algorithm respectively. As a result, the end products are predictable from time to time, and the time consumed to execute the algorithms are constant from time to time as well. Only through the operations of high precision nature, visual signals from different processing sources can be synchronized and integrated to the highest degree possible across the entire retina, which makes the entire visual field as a uniform whole.

To this end, the visual signals that originate from a lambda pair of cones in the fovea finally enter the four ganglion cell axons, two carrying the identical λ and two carrying the identical ι. There should be only one λ and one ι for the entire H cluster. They are the ultimate signals of color vision for the visual cortex to perceive the world in full color. Should the visual signals λ and ι carry the same sign as all positive or negative, they could confuse the visual cortex to mistake λ for ι or vice versa where they are combined and translated into colors.

4. One Dimensional Senses and Two Dimensional Senses
There are generally five senses in higher animals to collect information from the environment, vision, hearing, taste, smell, and touch. They can be further divided into two groups, one dimensional sense group or group one, which includes taste, smell, and touch, and two dimensional sense group or group two, which includes vision and hearing. The fundamental difference between these two groups is whether a single factor or two factors determine the nature of the neural signals that the central brain depends on to accurately perceive the external stimuli.

The external stimuli of the group one are usually chemicals that act on sensory cells that line the surface of a sensory organ to generate nerve impulses as responses. In the central brain the nerve impulses from the entire sensory organ are processed as a whole unit, not based on individual signals, to produce a single sense of the stimuli. The senses produced from kitchen vinegar are a good example to illustrate this. Vinegar is an aqueous solution of acetic acid. It displays a pungent smell to the olfactory cells and sour taste to the taste buds. Among human olfactory receptor family, one or more olfactory receptors can bind acetic acid molecules with different affinities and elicit neural responses of different degrees. The sense of smell will be stronger for receptors with higher affinity for acetic acid and weaker for receptors with lower affinities. The same happens to the taste buds. Nevertheless, the central brain will not distinguish which signals originate from receptors with higher affinity for acetic acid, and which signals are from the receptors with lower affinity for the acid. The smell of vinegar is the perception of aggregated presence of all the nerve signals arrived at the olfactory processing center from the nose without need to know the exact nature of individual nerve signals and the olfactory receptors that produced them, even though the degree of smell is proportional to the total number of olfactory receptors stimulated by the acetic acid molecules regardless of their binding affinities.

In general, perception of smell or taste or touch is the collective result of all the signals received from the entire peripheral sensory organ as a mixed whole without the need to decode each individual signals from the sensory cells. What matters is the concentration of chemical stimuli, the total number of stimulated cells, and the average affinities of all receptors for the chemical stimuli expressed on the sensory cells. More intense the average signals produced in the sensory organ, stronger the sense or perception generated in the brain, and higher sensitivities to the stimuli. Therefor, the sense of the group one stimuli is one dimensional, depending only on the collective amount of signals received, not on individual signals in the signal pool. In other words, the resolution of perception of smell, taste, and touch is on the level of the sensory organ, not individual signals and their producing cells.

One dimensional senses have evolved over a billion of years for the animals to perceive the presence of some chemicals or physical objects in the environment and respond accordingly. The sense of smell is especially critical to animals, especially to those that are solitary or have compromised vision, during mating seasons, hunting, danger detection, movement, familial member recognition, etc. It's obvious that in all these applications the overall sensitivity of the olfactory system matters most as animals rely on such a system to detect even a trace amount of chemicals like pheromones produced and released into the environment by other animals. Clearly this isn't dependent on the details of individual signals

among all responses from the olfactory system, but total amount of signals that reach the central brain from the nose.

Vision and hearing are group two senses that perceive physical entities, light and sound. A fundamental characteristic of the perception of physical entities is to form an image or sound in the central brain that exactly mirrors the external scene or sound, not just to generate a simple sense like smell being awful or taste being sweet. Perception of light or sound is actually the perception of color or sound wave, while color and sound wave are determined by the dual properties of wavelength and intensity. The intensity of group two is equivalent to the concentration of chemical stimuli in group one, however, every bit of the signal matters as it is the neurological representation of a particular point in the external visual scene or sound. Therefore, misrepresentation or deletion of one or more signals will corrupt or disrupt the entire scene or sound, thus forming an incomplete or distorted image or sound in the brain. Furthermore, all the signals arrived at the central brain must agree with the original scene or sound in spatial and temporal arrangement. In other words, any signals can't mix with other signals and must keep their own neurological identities in the peripheral and central brain. As a result, the amount of information for vision and hearing is far more overwhelming than for simple smell and taste. This raises a few basic questions. Are all-or-none universal action potentials sufficient to carry all the information for vision or hearing? Does each sensory cell have only one single signal transduction pathway to convey signals it produces to the central brain? Does the central brain require separate information about wavelength and intensity when perceiving color or sound? If yes, how and where are the light or sound signals decomposed into wavelength signals and intensity signals, in the sensory organ or central brain?

In mammalian auditory system, two types of hair cells in the inner ear are responsible for hearing. Most information about sound reached the brain via the inner hair cells, while the role of the outer hair cells is less clear despite over three times more numerous than the inner hair cells. Inner hair cells act as the mechanoreceptors for hearing: they convert sound frequency into electric activities and transmit them to the central nervous system. The function of the outer hair cells is now perceived as a sound amplifier by refining the sensitivity and frequency selectivity of the mechanical vibrations of the inner hair cells, thus likely determining the loudness in a frequency specific manner. Could the evolution of inner and outer hair cells be considered a coincidence like the presence of the dual channels in the visual system? The likeliness as a coincidence is small, but the electric signals for wavelength and intensity are indeed separated from each other for vision and hearing.

On the two dimensional coordinate system, let x axis be set for wavelength and y axis for intensity. Any point on the system defines the color or sound that corresponds exactly to one stimulus of the original signal source that falls on retina or inner ear, while all the points on the system constitute the entire scene or sound in a given moment of time. The positions of any points on the system are fixed relative to their neighboring points, and their neurological interpretation by the central brain can be different from their neighboring points as well. In other words, all the signals on the system constitutes an image of a visual scene or all sounds that enter eye or ear in a given moment of time. Such as coordinate systems must be present in the central brain in the visual cortex and auditory cortex as a mirror copy of the peripheral systems. A time series of such a coordinate system make up what we see or hear during the time covered by the time series, for example, a video or song clip. Therefore, vision and hearing are said to be two dimensional senses. In contrast, nerve signals for smell, taste, and pain can't be plotted onto a two dimensional coordinate system, but a one dimensional coordinate system.

Another fundamental difference between the two groups of senses is on the memory level. Human taste buds can distinguish many different tastes. If a drink contains honey, you will immediately recognize it

contains honey after a small sip of it. Humans can recognize a variety of food or chemicals based on their tastes in blind taste tests. However, can you project the honey taste in your brain when you are thinking of honey? Similarly can you project the smell of vinegar or a certain fragrance, or the pain of a prickle on the skin in your brain when you are thinking of them? The answer is a clear no. Human brains are unable to generate the senses of smells, tastes and touch in the central brain in the absence of chemical or physical stimulation of the sensory organs. A possible explanation is that group one senses are hard to define precisely in terms of nerve signals as they are treated as a whole, not individual signals, posing difficulty of how and what portion of the signals to store. What the central brain can do best seems to store a minimum amount of information that is just enough to reminder the brain that the incoming signals are from olfactory receptors or taste buds or pain sensors, and what smell, taste, or pain they are.

On the other hand, an image or sound can appear in human brain in the absence of stimulation of the real object or sound that appears in the brain. For example, you can project an image of rose or play a song in your brain. In this case, the image or sound must be stored in the memory in early experiences.

Human thinking is actually the result of the silent sound in the form of human language endlessly going throughout a thinking process. More importantly, the sound is the replay of a tone in the brain that the individual knows. For example, the tone of someone's voice, a particular car's engine, a bird song, etc. The sound in the replay is silent and flat, meaning the loudness of the sound can't be raised or lowered. The tone of a sound is determined by the sound wavelength, suggesting that the wavelength from a given sound source is stored in the memory, while the loudness or amplitude of any sound source isn't. It makes sense that it is unnecessary to record the loudness of a sound as the loudness is a general characteristic of sound applicable to all the sounds. An implication of the sound replay is that the sound signals originated in the ear are separate in two types in the central brain, one encoding wavelength for the tone and one encoding amplitude for loudness. Only wavelength part of the signals forms memory, while the amplitude is lost after its use. An immediate question is that where the nerve signals for the sound are separated into two types, in the ears or in the central brain? A logic answer is in the ear as the ear has the first hand information about the physical dual properties of the sound. The presence of inner hair cells and outer hair cells seems to be a must for bringing sound into the central brain not only for perception, but also for memory storage. If evolution didn't make inner and outer hair cells, it would make other cellular structures to fulfill the same functionalities.

The similar situation is less clear for vision. Like hearing, the presence of the dual channels and super ribbon triad structures in one channel isn't a pure coincidence as well, but a must for bringing light into the central brain for perception. However, is it necessary for memory storage? It seems uncertain which visual signals, wavelength or intensity or both is stored in memory for future image projection. Unlike sound replay in which tones are clear, accurate, and flowing unbroken for easy recognition, and even enjoyment, the image projected from memory is unclear and distant with poorly defined boundary and shapes. It requires effort of mental concentration to project as the image will disappear quickly after concentration is eased. Can you project color red in your brain? Does the flower rose projected from the memory have the color of dark red like a natural rose the eye sees? The projected image is dull regardless even if it is in color. Unlike sound amplitude which is a general factor to alter the level of the tone, the intensity of light is one of the two deterministic factors for the color as shown in Figure 3. An immediate question is that which visual signals from the memory, λ or i or both, are used to project images in the brain? Or which visual signals, λ or i or both is stored in memory?

How an image is created? By observing maple tree leaves from spring to fall, we found that all leaves remain in the same shapes, but their colors are changing as seasons change. The surface of a leaf

reflects the light into our eyes, resulting in a leaf shaped area on the retina being stimulated and then generating an image of that leaf in our brain. While the same leaf reflects light into our eyes, always forming the same shaped leaf image, color changes in different reasons cause leaves to reflect light of different wavelength into our eyes, resulting in the same leaf image with different colors. It can be concluded that it's the quantity of photons or intensity, not wavelength of the photons, that delineates the shape of an image. In other words, it's the area in the retina that is hit with photons defines the bounds of an image, while wavelength is only an attribute of light that gives the image colors. The same wavelength will confer any shaped surfaces the same color as the same paint will paint any surface to the same color.

Regardless, memory based image projection and the anatomical structure of retina substantiate the conclusion that the visual signals have diverged into separate signals for wavelength and intensity along the signal transduction pathway prior to their entry into the central brain. And the two types of signals are treated differently in terms of memory storage. The wavelength determines the base color, while intensity determines the brightness of that base color in quite dramatic way. The combination of wavelength and intensity generates millions of colors human eyes can distinguish. Nevertheless, it's the intensity that finally determines the shapes of a physical objects in addition to color, and it's logical to think that wavelength and intensity shouldn't be stored as single signals, but separate signals for the color. It is certain that ɪ is the visual signals for physical objects that are stored into memory and it's also the memory source to project images of the physical objects. The two dimensional senses are more than generating an overwhelming amount of information for best perception, they also expose another layer of mystery how the central brain processes and store information gathered from the surroundings through vision and hearing.

Nerve signals from the one dimensional peripheral sensory organs are stored in memory as concept memory, while for nerve signals derived from the two dimensional peripheral sensory organs, one dimension is stored as concept memory and the other as physical memory. Only physical memory can be retrieved to project the original senses. Earlier discussion indicates that the nerve signals that represent light intensity and sound wavelength are stored as physical memory, while the nerve signals representing light wavelength are stored as concept memory. Nerve signals for sound amplitude may not be stored at all.

Existence of concept and physical memories reveals different memory mechanisms at work in the central brain. If we dig a little deeper, we can ask a very basic question how many memory neurons are engaged in memorizing the pungent smell or sour taste of vinegar, or the color red? If you looked at a small red circle and a large red square, your brain formed memory for the shapes of circle as well as square, but how could your brain form memory for the color red? Color is like smell or taste or pain except that color can't exist on its only, but always attached to a physical object. The red on the circle and square are the same, so the nerve signals that represent them. Should the nerve signals of different origins for the color red be stored in separate memory spaces in the memory? If this was the case, then color red would be stored in memory millions of time after we looked at red color millions of time. For concept memory, therefore, a small number of neurons suffice to mark the smell of vinegar or color red, and the nerve signals from identical stimuli like color red or vinegar will be stored only once. An advantage of the concept memory is store once and use everywhere with the same smell, taste, pain or color, a clear approach to save memory neurons and avoid redundant storage. A critical issue is that the memory must store concept information enough to distinguish one concept from vast other concepts stored in memory. In contrast, physical memory stores a concrete physical entity which has bounds and details, thus requiring more neurons to store.

The purpose of this section is to provide additional neurophysiological evidence to substantiate the conclusion that the raw visual signals separate into λ and i, not only for the best accuracy and acuity, but also for memory storage.

5. Conclusions and Discussion

This paper is intended to explain how the color vision works with an emphasis on the central vision. Visible light possesses two independent physical properties, wavelength and intensity. In color theory, hue, along with saturation and brightness, makes up the three distinct attributes of color. Hue is related to wavelength of visible light, and the brightness depends upon the amount of light falling on a surface, which is proportional to the amount of light emitted from a given color source. Saturation pertains the amount of white light mixed into a hue. A fully saturated color is any spectral color that consists of only one wavelength. Saturation reflects the degree of how mixed the wavelengths of light are. In a sense, both hue and saturation are related only to the wavelength of light. The composition of wavelengths is intrinsic to the light emitted from given sources, but the intensity can vary as the power of light sources vary. Because color is perceived by the eye, the dual properties of light must be the two fundamental factors the eye relies on to achieve color vision. Therefore, the cellular and molecular basis for color vision is predicted to generate the neural forms of wavelength and intensity for the central brain that represent the physical wavelength and intensity, respectively.

Retina is such a fascinating masterpiece created by mother nature to perceive the world that emits full spectra of light. The simple, raw visual signals start from light exposure in the photoreceptors and undergo a long and baffling journey to finally end up in the central brain. In this journey the retina does all the nuts-and-bolts kind of signal processing jobs with only one goal: to reduce the primitive visual signals – electric membrane potential $\mathcal{V}$s – into neural wavelength and intensity needed for perception of color in the visual cortex. Let's look at the Equations (1) to (4) again in one place:

$$\mathcal{V} = \mathbf{R}\mathbf{I} \tag{1}$$

$$\lambda = \frac{\mathbf{R}_1}{\mathbf{R}_2} = \frac{\mathcal{V}_1}{\mathcal{V}_2} \tag{2}$$

$$\mathcal{V}_\lambda = \frac{\mathcal{V}}{\Delta\mathbf{R}} \tag{3}$$

$$i = \mathcal{V}_{\lambda 1} - \mathcal{V}_{\lambda 2} \tag{4}$$

Equations (1) to (4) lay the architectural blueprint of the cellular structures for color vision, and the primate retina is the biological implementation of Equations (1) to (4). Equation (1) requires two cell types with different sensitivities to the same wavelength to produce a pair of electric membrane potentials – lambda $\mathcal{V}$s. In primates, M-cone/L-cone pair and S-cone/M-cone pair fulfill Equation (1). Equations (2) and (3)/(4) dictate the generation of λ and i in separate spaces and neuronal operations, respectively. In primates there exist two channels, N channel for processing and conducting λ signals, and P channel for processing and conducting i signals. Equation (3) requires that the initial electric membrane potentials $\mathcal{V}$s be modified in certain way for Equation (4). In primates, $\mathcal{V}$s are modulated first into $\mathcal{V}_\lambda$s in the super cellular structure ribbon triads before they are used as input for i production in sub-lamina b. Upon implementation of Equations (1) to (4), the primate retina has achieved the capability of full color perception. It first converts light into raw electric membrane potential $\mathcal{V}$s in the cones and then reduces $\mathcal{V}$s back into neural forms of the wavelength λ and intensity i via two signal

transduction pathways composed of bipolar cells, ganglion cells, and accessory horizontal cells and amacrine cells. λ and $\mathfrak{i}$ are the final input to the central visual cortex for color vision.

When light enters the eye, it strikes individual cones to produce electric membrane potentials, the initial visual signals $\boldsymbol{\nu}$s. $\boldsymbol{\nu}$s are the sole raw signals left by light stimulation. Because of different sensitivities of the visual pigments in the cone cells to the same light, production of $\boldsymbol{\nu}$s in cones is not sufficient for color vision as $\boldsymbol{\nu}$s aren't the truthful reflection of the light that entered the eye. Therefore, retina must de-correlate $\boldsymbol{\nu}$s into wavelength and intensity by transforming $\boldsymbol{\nu}$s into signals that truthfully represent light free of bias by removing the effect of cone spectral absorption variations.

According to Equation (1), a single cone type can be excited by light of different combinations of wavelength and intensity to produce the same amount of $\boldsymbol{\nu}$s, making it insufficient to differentiate a response change in wavelength from a response change in intensity. The visual information must be obtained by comparing the responses elicited by the same light in two or more types of cones. The complexity of the visual pathway is evolved to work around the principle expressed in Equation (1), and any theories in color vision field should at least be able to give plausible accounts of how different types of cones cooperate to achieve spectral decorrelation.

Equations (1) to (4) not only lay the cellular organizational blueprint for color vision, but also equally importantly they establish how physical wavelength and intensity of light will be represented in neural forms free of bias. Based on Equation (1), the initial electric potential $\boldsymbol{\nu}$s are linear and constant to a given source of light, and the ratio of the lambda $\boldsymbol{\nu}$s is a constant, dependent only on wavelength (Equation 2). This can't be a coincidence as its biological significance is apparent. It has the quality to be the neural form of wavelength or λ that represents the physical wavelength of light for the central brain to perceive hue. On the other hand, the difference of lambda $\boldsymbol{\nu}$s is less sensitive to cone's spectral absorption variations and can be used to derive the neural form of intensity that represents the physical intensity for the central brain to perceive brightness of light. However, the neural form of intensity obtained in this way isn't the best to represent the true intensity of light, since the effects of the same wavelength on $\boldsymbol{\nu}$s are not the same in different cone types. The effect of wavelength on lambda $\boldsymbol{\nu}$s is called lambda effect. The retina needs to resolve lambda effect on each $\boldsymbol{\nu}$ of the lambda pair before taking their difference. This difference is the neural form of intensity or $\mathfrak{i}$, representing the physical intensity of light as shown in Equation (4). Resolution of lambda effect is the extra step predicted by Equation (3) that the retina must take to produce $\mathfrak{i}$.

In retina, a single cone passes its signals to two bipolar cells. One bipolar cell hyperpolarizes on receiving the signals from the cone, thus negative or N bipolar cell. This N bipolar cell passes its signal to an N ganglion cell in the same fashion. A pair of N bipolar cell and N ganglion cell makes up the negative or N channel. Another bipolar cell depolarizes on receiving the signals from the cone, thus positive or P bipolar cell. This P bipolar cell passes its signals to a P ganglion cell in the same fashion. A pair of P bipolar cell and P ganglion cell makes up the positive or P channel. The cellular structural evidence for the two channels in the retina strongly suggests that visual signals are partitioned into two channels, upholding the presence of the neural forms of wavelength and intensity as predicted by the Equations (2) and (4). Such a dual channel system guarantees that their production doesn't interfere with each other in space and time. At this point there is no clue which channel is for which form of the neural signal, λ or $\mathfrak{i}$.

Equation (3) predicts that there must exist a special cellular structure in one of the channels to resolve lambda effect, and the channel with such a structure is for production of the neural form of intensity $\mathfrak{i}$. A super structure, called ribbon triad, is located at the bottoms of cone axon terminals and controls the

gate of the P channel, indicating that it is the site where the lambda effect is resolved and the P channel has been evolved for the i. The ribbon triad consists of a cone axon terminal, a dendritic tip of a P bipolar cell and two dendritic tips from two horizontal cells. The triad structure is also predicted to be the site where cones across the fovea are matched into lambda pairs. The triad structures make sure that the i values are produced to the possible maximum accuracy for the best color resolution and acuity. Are the presence of the ribbon triad structure and resolution of lambda effect a pure coincidence? No, it isn't. If it was, what would be the roles for the triads in visual signal processing? At the triads, primitive visual signals lambda $\boldsymbol{\nu}$s are modulated to become lambda $\boldsymbol{\nu}_\lambda$s at the right time just before entering the P bipolar cells.

Sandwich algorithm is a hypothesized working model to identify lambda pairs. Lambda pairs that share the same set of $\boldsymbol{\nu}$s fuse into an H cluster, forming a group of cone cells, which are heterogeneous in cone types, but homogeneous in light exposed. A visual scene is delineated into a number of H clusters, depending on its complexity.

Further processing of lambda $\boldsymbol{\nu}$s and lambda $\boldsymbol{\nu}_\lambda$s in their own channel generates λ and i, the final formats for the central brain. To this end, all ganglion cells descending from the cone cells in an H cluster carry the same λ in the λ or N channel and the same i in the i or P channel. Bipolar cells and ganglion cells seemed to provide a signal passage that allowed horizontal cells and amacrine cells to transform signals into the final formats as they are traveling down through it.

Regardless of sick or healthy, happy or sad or mad, young or old, in cold or hot weather, the perception of the same natural scene is never changed if the scene is not changed physically. The retina is truly a biological computer with extraordinary precision, stability, repeatability, and efficiency. Like a modern computer, the biological computer must rely on some neural algorithms to guarantee that all nervous signals are processed, coordinated, integrated, synchronized, and transmitted precisely, repeatably, and efficiently in input, output, space, and time. At least our neural algorithms – resolve algorithm, match algorithm, λ algorithm, and i algorithm – are working to guarantee the smooth transformation of visual signals. With the aid of computer simulation technology, it would be possible to simulate some of these neural algorithms to reveal the true nature of color vision.

A neural algorithm must be supported by special cellular structures, and a special cellular structure may be evolved to bring a particular neural algorithm into being. The cellular structure could provide a hint of how a neural algorithm would work. The four neural algorithms do show certain structure-function correlations with the cellular structures in the retina where these algorithms are executed. The resolve algorithm is intended to resolve the lambda effect on early visual signals to make the i values closer to the physical intensity of light. It may incorporate additional information from nearby cones to have a broader evaluation of the signals for better resolution. The underlying super ribbon triad structure forms a special micro-environment in which $\boldsymbol{\nu}$s and additional information received from nearby cones via horizontal cells interact and influence each other to change the amount of glutamate released into the synaptic cleft, thus changing the $\boldsymbol{\nu}$s for the P bipolar cells. The triad structure must maintain a set of neurotransmitter receptors, ion transporters, ion channels, etc with strict on/off threshold values to control the release of glutamate according to the algorithm. Match algorithm is a sorting algorithm acting on multiple cone cells to group them into lambda pairs. Match algorithm relies on the triad as its underlying cellular structure to collect $\boldsymbol{\nu}$s from an area covered by its horizontal cells. Both λ and i algorithms merge lambda $\boldsymbol{\nu}$s and $\boldsymbol{\nu}_\lambda$s to generate new $\boldsymbol{\nu}$s – λ and i. The underlying cellular structures are dense axon-dendrite connections, and the connections are heavily intertwined with amacrine cells, the accessory cells between OPL and IPL. These structures can support extensive communication between individuals of the lambda pair, partition signals along the axon and dendrite tree terminals for

micromanagement, and provide the means to exchange information with local and global bipolar and ganglion cells for coordination, integration, synchronization, and so on.

Separation of visual signals into λ and ι have great advantages in signal processing. If there was only a single channel, for example, the widely accepted ON channel (see below) to process, carry, and convey visual signals, would the eye be able to distinguish a million of colors, most of which differ only subtly or barely detectably? The answer is a definitive no because the all-or-none action potentials are too narrow in their range to distinct signals which differ in all degrees, tiny or big. However, separation of signals into λ and ι confers visual system a combinatorial power to double the capacity to produce and conduct visual signals in more granulate details. In addition, the dual channel system allows production of ι in its own processing space, thus broadening the ι range without impact on λ, and enhancing the visual sensitivity to tiny changes in intensity. Lastly because λ is chromatic and ι is achromatic, the central brain can perceive colors by adding chromatic λ to achromatic ι much like adjusting achromatic paint to desired color tone by adding more or less chromatic solvent. Color comparison is made more easily and precisely by comparing λ with λ, and ι with ι.

Physics of visible light, cellular structures of the fovea area in the retina, the necessity of multiple types of cones for color vision are all hinting a scheme of visual information flow that starts from the cones and ends at optical nerve described by Equations (1) to (4). A theory in color vision must address the possible roles and action of each of the cellular structures and components present in the visual system and provide reasonable account for how different types of cones work to resolve ν changes caused by wavelength or intensity, and how weakly absorbed red light is perceived as brightly and vividly as strongly absorbed green or yellow light. Otherwise it isn't a theory per se and has no value in science.

In a logical design of workflow, each step is given a specific task to accomplish. The later step will continue what's left by the earlier steps and won't repeat the tasks completed by the earlier steps unless there is a need to repeat. All living organisms are highly efficient in utilizing the resources available to them, and the retina won't be exception. Visual pathway isn't long, consisting of a cascade of only three cells of different types, which enables the visual signals to be processed in a highly sequential and logical order along the pathway. It's expected that each cell type has its own special roles in this logic workflow that other cell types don't have, ensuring that the signal processing is a stepwise process with each step accomplishing a specific task, and that the later cell types in the pathway won't be possibly performing any tasks already performed by earlier cell types.

A logic order of visual signal processing would be likely in an order of local scope to global scope. The earliest step generates the raw νs upon light exposure. When the cone cells pass νs to bipolar cells, early transformation happens, but it is limited to the local scope just to do what early step can do in the pathway. It would be illogical to think that the raw νs would be modified in an all-around fashion in this early time before other relevant information was available from wider areas, as this could bear serious consequences of distorting the color. A logical work in this transformation is to partition the raw νs into N bipolar cell and P bipolar cell, respectively, making ready for the generation of λ and ι in the next step in the pathway. Partition of the raw νs must be done in the early time in this workflow, as delay will lead to signal contamination and loss of signal identity. Resolution of lambda effect must be done here as ι production must be in sync with λ production. This early transformation doesn't require collection and integration of information from cones located far away.

After the raw νs are partitioned into N and P bipolar cells, transformation of νs into λ and ι follows immediately since production of λ and ι must depend only on the νs directly from the cone cells to ensure that neural forms of wavelength and intensity are genuinely reflecting those of the light. This is

especially essential for **λ** as **λ** is the base color which will paint the **ɩ** to generate color. Therefore, production of **λ** and **ɩ** is local as well.

Signal transformation is local and ends upon the completion of production of **λ** and **ɩ**. Only when **λ** and **ɩ** pass down through the later part of the logic workflow in their own channel, it makes sense for the neural form of intensity to subject to further modulation in a global scope. Global scope means the need to collect processed information from a large number of bipolar cells and ganglion cells from a broad area and use this information to evaluate and then make necessary modulation to the **ɩ** values so that the **ɩ** fits well in the visual scene. In the case of possible contrast enhancement, which obviously is a global action, the retina must first know what visual signals are involved in contrast and when they are available for comparison in order to adjust and enhance the contrast. Therefore, only when the global scope signals are available in the later steps of the logic workflow, signal modulations like contrast enhancement can occur if they are part of the visual processing pathway.

Color perception under the same lighting condition is consistent in any moment of time. Because of this, a coefficient can be proposed to describe the phototransduction pathway and express a quantitative relationship between generation of transmembrane electric potentials, photon absorption rate, and the number of photons hitting a cone in a unit of time as Equation (1). Equation (1) is the base equation from which Equations (2) to (4) are derived. Equations (1) to (4) establish a theoretical logic workflow for visual signal processing, which agrees well with what has been discussed earlier, which is based on observations on molecular and cellular structures of the retina and electrophysiological study conducted during light stimulation. It's particularly interesting that Equation (3) predicts the existence of an extra cellular structure in the visual pathway required for **ɩ** generation. All this can't be a coincidence, but is what exactly happens in the visual perception of color.

The processed visual signals **λ** and **ɩ** converge into the optical nerve where they are dispatched to the visual cortex in the central brain. When nervous impulses enter the central brain from the peripheral, they have entered a totally different world, where they all have lost their identity and individuality like a drop of water mixed into a pond. They form a whole with all the other nervous impulses from the peripheral. In the central brain, the whole is much much greater than the sum of its parts, resulting in the impossibility to trace the origin of each of the individual impulses in the whole to the peripheral. It must be cautious in attempt to establish what happens in the peripheral by extrapolating from what we all are experiencing in the central brain, as the conclusions or results are more likely to misguide and, as a result, hinder research.

N bipolar cells are called OFF center bipolar cells and P bipolar cells ON center bipolar cells in the research literature, and same for ganglion cells. ON center bipolar cell and ON center ganglion cells make up the ON center channel, and OFF center bipolar cell and OFF center ganglion cells make up the OFF center channel. Furthermore, Off center cells and ON center cells form a center surround organization in a configuration of color opponent centers and surrounds (red/green) and ON or OFF centers and surrounds. The concept ON/OFF center surrounds and color opponent center surrounds is too overly complicated and abstract to be true in retina. It's hard to apprehend what possible concrete roles the center surround organizations and ON/OFF center channels could serve in color vision, especially in terms of wavelength and intensity, super ribbon triad structure, L-cone/M-cone pairs for color resolution, and perception of spectrtally weak light. Why does a cone diverge its signals into ON and OFF center bipolar cells? What nature are the signals in the ON and OFF center channels to color vision? In individuals with normal color vision the ratio of L-cone to M-cone can be off 1:1 mark considerably to 10:1 or 1:10, or worse. An inference is that a cone of one type can't be surrounded by cones of other types in individuals with off-mark ratios. The rigidity of the center surround organization

is also at odds with the unlimited capacity to respond to all natural scenes from surroundings. The center surround organization is established largely on data obtained from intracellular and extracellular recordings of bipolar and ganglion cells, hence its validness is highly questionable. However, these theories have become the foundation of color vision research. Research based on wrong theories is detrimental because it can misguide experimentation, thinking, and interpretation, all of which can impede the progress in our understanding of color vision.
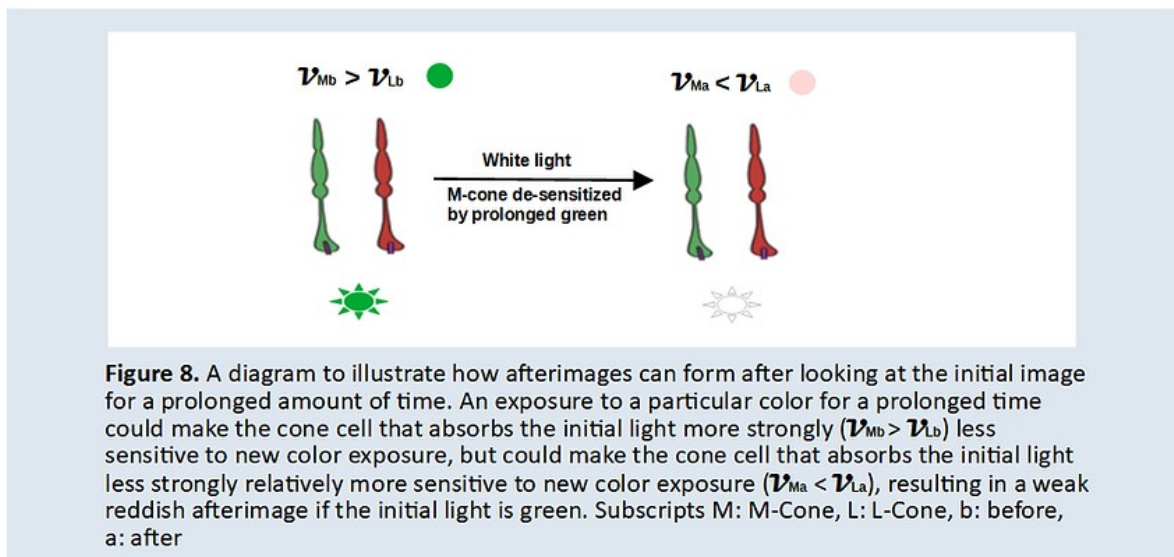
More importantly in phototransduction, light stimulation turns photoreceptors hyperpolarized. Are light stimulated photoreceptors OFF? Hyperpolarization is simply a neural state after a neuron is exposed to certain external stimuli. As long as an hyperpolarized state could impact downstream neurons, it should not be considered OFF. Furthermore, an OFF neuron could turn the postsynaptic cells ON if receptors expressed on the postsynaptic cells could lead the cells to depolarize upon receiving the pre-synaptic impact. At least in the visual system, the sign of the responses from external or internal stimuli can't be the determinant factor for responses to be on or off. Calling N bipolar cells OFF and P bipolar cells ON is misleading and improper.

To end this paper, let's use λ values to explain how the colors of multi-wavelength light are determined. Naturally light that enters the eye is not of single wavelength, but of mixed wavelengths. The 11-cis retinal-opsin molecule would absorb any photons in its absorption range. Each wavelength in multi-wavelength light had its own intensity, and stimulated cone cells independent of other wavelengths in the same light. The end result was another $\mathcal{V}$, the sum of all voltage changes evoked in the cone by each wavelength. That meant the λ would be generated from the lambda $\mathcal{V}$s derived from mixed-wavelength: $\mathcal{V}1/\mathcal{V}2$ or $R1/R2$, where $\mathcal{V}1$ and $\mathcal{V}2$ are composite $\mathcal{V}$s, and $R1$ and $R2$ are composite constants. If this λ happened to be equal to the λ generated from the lambda $\mathcal{V}$s derived from single-wavelength light, the color would be perceived as that single-wavelength color. Because cones discriminate wavelengths only through its retinal-opsin complexes via different absorbance spectra, it would be impossible for the retina to know whether the $\mathcal{V}$, so the λ, was the result of exposure to single wavelength light or multi-wavelength light. Rather the retina only recognizes the signals it generates, and doesn't care how these signals are generated and if λ and i are composite values or not.

If light contains two wavelengths that are close in the visible spectrum, like half green and half yellow, the total $\mathcal{V}$t would be the sum of green $\mathcal{V}$ and yellow $\mathcal{V}$. The size of $\mathcal{V}$t fell between the two $\mathcal{V}$s if the light contains pure green and pure yellow respectively. The λ value computed from the lambda $\mathcal{V}$ts fell just within the space close to both green λ and yellow λ, which is an additive color – greenish yellow. On the other hand, if light contains two wavelengths that are not close in the visible spectrum, like half green and half red, the total $\mathcal{V}$t would still be the sum of green $\mathcal{V}$ and red $\mathcal{V}$. The size of $\mathcal{V}$t again fell similarly between the two $\mathcal{V}$s if the light contains pure green and pure red respectively. Because the difference between green $\mathcal{V}$ and red $\mathcal{V}$ was quite large, $\mathcal{V}$t would be close to neither green $\mathcal{V}$ nor red $\mathcal{V}$, and the λ value computed from the lambda $\mathcal{V}$ts fell within a space between green λ and red λ, but close neither to green λ nor red λ. As a result, mixing green and red wouldn't lead to an additive color – greenish red, but a color between green and red, that was yellow. A similar picture could be drawn for light containing blue and yellow. Yellow and green seem to be two chasms in the visible spectrum, and mixing wavelengths from the two sides of each chasm generates colors that fall within that chasm. This is obvious from the visible spectrum. From color opponency point of view, it seems that wavelengths from the two sides of each chasm opposes each other, but actually, it is simply the fact that mixing two dissimilar things lead to a thing that is dissimilar to both of its parents, but stands on its own. There is no such thing as color opponency.

Finally an attempt is given to explain how afterimages are produced. The central theme of the dual channel system is that the color we perceive through eye is deduced from the sole products of phototransduction – the electric transmembrane potential $\mathcal{V}$s – that occurs in M-cones and L-cones. Any visual phenomena must be explained by focusing on what normal visual processes have been disrupted in the retina that caused the visual phenomena in question. If this is not the case, then the cause could be residing in the visual cortex.

Figure 8 illustrates how an afterimage could form after a prolonged exposure to green light. When exposing to green light, M-cone, which absorbs green light more strongly than L-cone, produces a membrane potential $\mathcal{V}_M$, which is greater than the membrane potential $\mathcal{V}_L$ produced by L-cone. Upon prolonged exposure to green light, M-cone becomes less sensitive to new color exposure, in this case, the white light. This makes L-cone relatively more sensitive to the white light. As a result, the sizes of the membrane potentials are reversed from $\mathcal{V}_{Mb} > \mathcal{V}_{Lb}$ to $\mathcal{V}_{Ma} < \mathcal{V}_{La}$, here subscript b means before, and a means after. The direct consequence of this reverse is temporarily shifting λ from green zone into the weak reddish zone, which is perceived as a weak reddish afterimage.

Figure 8. A diagram to illustrate how afterimages can form after looking at the initial image for a prolonged amount of time. An exposure to a particular color for a prolonged time could make the cone cell that absorbs the initial light more strongly ($\mathcal{V}_{Mb} > \mathcal{V}_{Lb}$) less sensitive to new color exposure, but could make the cone cell that absorbs the initial light less strongly relatively more sensitive to new color exposure ($\mathcal{V}_{Ma} < \mathcal{V}_{La}$), resulting in a weak reddish afterimage if the initial light is green. Subscripts M: M-Cone, L: L-Cone, b: before, a: after

Changes in the membrane potential $\mathcal{V}$s in cone cells could result in the changes in neural λ, which changes our perception of color, which underpins the notion that the dual channel system of the retina is at work for the color vision. It must be noted that it can have many plausible explanations for a neural phenomenon because the whole is much greater than the sum of its parts for the nerve system. It doesn't make sense to fix only on one explanation, while ignoring all other possibilities, even better possibilities.

Send correspondence to Hangjiong Chen, PhD, at hjchen1@yahoo.com