A Bio-inspired Synergistic Virtual Retina Model for Tone Mapping

Marco Benzi, Maria-Jose Escobar, Pierre Kornprobst

A Bio-inspired Synergistic Virtual Retina Model for Tone Mapping

(https://doi.org/10.1016/j.cviu.2017.11.013)

Marco Benzi[a,b], María-José Escobar*[a], Pierre Kornprobst*[b]

* *Co-senior-authors*

[a] *Universidad Técnica Federico Santa María, Departamento de Electrónica, 2390123 Valparaíso, Chile*
[b] *Université Côte d'Azur, Inria, France*

Abstract

Real-world radiance values span several orders of magnitudes which have to be processed by artificial systems in order to capture visual scenes with a high visual sensitivity. Interestingly, it has been found that similar processing happens in biological systems, starting at the retina level. So our motivation in this paper is to develop a new video tone mapping operator (TMO) based on a synergistic model of the retina. We start from the so-called Virtual Retina model, which has been developed in computational neuroscience. We show how to enrich this model with new features to use it as a TMO, such as color management, luminance adaptation at photoreceptor level and readout from a heterogeneous population activity. Our method works for video but can also be applied to static images (by repeating images in time). It has been carefully evaluated on standard benchmarks in the static case, giving comparable results to the state-of-the-art using default parameters, while offering user control for finer tuning. Results on HDR videos are also promising, specifically w.r.t. temporal luminance coherency. As a whole, this paper shows a promising way to address computational photography challenges by exploiting the current research in neuroscience about retina processing.

Keywords: Tone mapping, HDR, retina, photoreceptor adaptation, contrast gain control, synergistic model

1. Introduction

Real-world radiance values span several orders of magnitudes which have to be processed by artificial systems in order to capture visual scenes with a high visual sensitivity. Think about scenes of twilight, day sunlight, the stars at night or the interiors of a house. To capture these scenes, one needs cameras capable of capturing so-called high dynamic range (HDR) images, which are expensive, or via the method proposed by [21], currently implemented in most standard cameras. The problem is how to visualize these images afterwards since standard monitors have a low dynamic range (LDR). Two kinds of solutions exist. The first is technical: there are HDR displays, but they are not affordable for the general public yet. The second is algorithmic: there are methods to compress the range of intensities from HDR to LDR. These methods are called tone mapping operators (TMOs) [84]. TMOs have been developed for both static scenes and videos rather independently. There has been intensive work on static images (see [49, 11] for reviews), with approaches combining luminance adaptation and local contrast enhancement sometimes closely inspired from retinal principles, as in [64, 8, 30, 67] just to cite a few. Recent developments concern video-tone mapping, where a few approaches have been developed so far (see [28, 27] for surveys).

Interestingly, in neuroscience it has been found that a similar tone mapping processing occurs in the retina through adaptation mechanisms. This is crucial since the retina must maintain high contrast sensitivity over this very broad range of luminance in natural scenes [86, 34, 44]. Adaptation is both global through neuromodulatory feedback loops and local through adaptive gain control mechanisms so that retinal networks can be adapted to the whole scene luminance level while maintaining high contrast sensitivity in different regions of the image, despite their considerable differences in luminance (see [89, 93] for reviews). Luminance

and contrast adaptation occurs at different levels, e.g., at the photoreceptor level where sensitivity is a function of the recent mean intensity, and at the bipolar level where slow and fast contrast adaptation mechanisms are found. These multiple adaptational mechanisms act together, with lighting conditions dictating which mechanisms dominate.

Thus there is a functional analogy between artificial and biological systems: Both target the task of dealing with HDR content. Our motivation is to develop a new TMO based on a synergistic model of the retina. We start from the so-called VIRTUAL RETINA simulator [101] which has been developed in neuroscience to model the main layers and cell types found in primate retina. It is grounded in biology and it has been validated on neurophysiological data. As such, it has been in several theoretical studies as a retina simulator to make predictions of neural response [61, 7, 96, 62, 13]. Thus this model can be considered as a good candidate to build a synergistic model since it could pave a way for much needed interaction between the neuroscience and computer vision communities, as discussed in the survey by [63].

Interestingly, VIRTUAL RETINA has also been used to solve artificial vision tasks, such as hand written recognition [66] and image compression [60, 22, 23]. The reason why it could be also an interesting model for a TMO is that it includes a non-trivial Contrast Gain Control (CGC) mechanism, which not only enhances local contrast in the images but also adds temporal luminance coherence. However, VIRTUAL RETINA is missing several important features to address the tone mapping task. It was not designed to deal with color images, even less HDR images, there is no photoreceptor adaptation nor readout mechanisms to create a single output from the multiple retinal response given by different output cell layers. In this paper we address these questions by enriching VIRTUAL RETINA model keeping its bio-plausibility and testing it on standard tone mapping images.

This article is organized as follows. In Sec. 2 we describe our bio-inspired synergistic model and we compare it to former bio-inspired TMOs In Sec. 3 our method is evaluated on static images and videos. We show the impact of the different steps and of main parameters. We show comparisons with the state-of-the-art. Finally, In Sect. 4 we summarize our main results and discuss future work.

2. Bio-inspired synergistic retina model

2.1. Why VIRTUAL RETINA*?*

In this paper we are interested in investigating how the retina could be a source of inspiration to develop a new TMO. In neuroscience, there is no unique model of the retina but different classes of mathematical models depending on their use. In [63], the authors identified three classes of models: (i) linear-nonlinear-poisson models used to fit single cell recordings [16, 12], (ii) front-ends retina-like models for computer vision tasks [9, 39] and (iii) models based on detailed retinal circuitry [101, 54]. There is no unique model also because the retinal code remains an open challenge in neuroscience. With the advent of new techniques such as Multi-Electrode Arrays, the recording of the simultaneous activity of groups of neurons provides a critical database to unravel the role of specific neural assemblies in spike coding.

There is currently an intensive ongoing research activity on the retina trying to decipher how it encodes the incoming light by understanding its complex circuitry [35]. As such, models based on detailed retinal circuitry will directly benefit from this research and this is why it seems promising to focus on this class of models. Here, we consider the VIRTUAL RETINA model [101] as starting point. VIRTUAL RETINA was designed to be a spiking retina model which enables large scale simulations (up to 100.000 neurons) in reasonable processing times while keeping biological plausibility. The underlying model includes a non-separable spatiotemporal linear model of filtering in the Outer Plexiform Layer, a shunting feedback at the level of bipolar cells, and a spike generation process using noisy leaky integrate-and-fire neurons to model RGCs. To be self contained, this paper will remind the main equations of VIRTUAL RETINA but we will not enter their justifications, so the interested reader should refer to the original paper.

2.2. Model overview

Our TMO has three main stages shown in Fig. 1:

(i) Pre-processing steps including a conversion of HDR-RGB input color image to luminance space, and a calibration of this luminance data so that these values are mapped to the absolute luminance of the real world scene.

(ii) A detailed retina model including a new photoreceptor adaptation taking into account pupil adaptation, which has been incorporated into the input level of VIRTUAL RETINA model that follows.

(iii) Post-processing including the definition of a readout given the multiple retinal output of VIRTUAL RETINA and followed by colorization and gamma correction to finally obtain a LDR-RGB output image.

Each stage is explained in detail in the following section and also illustrated using a representative example to show its impact. The source code is available upon request to the authors via the paper website.[1] Website contains sample results for static images and videos, and the download form for the source code.

2.3. Model detailed description

Input is a HDR video defined by $(L_{\text{red}}, L_{\text{green}}, L_{\text{blue}})(x, y, t)$, where L_{red}, L_{green} and L_{blue} represent the red, green and blue channels of the input HDR video respectively. Variables (x, y) stand for spatial position and t for time.

2.3.1. Pre-processing: Color-to-luminance and HDR calibration

This first stage is classical. One first convert HDR color video into a HDR radiance map video following the linear operation:

$$L_{\text{w}}(x, y, t) = 0.2126\ L_{\text{red}}(x, y, t) + 0.7152\ L_{\text{green}}(x, y, t) + 0.0722\ L_{\text{blue}}(x, y, t).$$

This equation correlates the way light contributes differently to the retina depending on the proportion of each photoreceptor type with green the most energy and blue the least. Here we use the standard matrix specified by ITU-R BT.709 for the input channel weights [84], also used in the sRGB color space.[2] This is the most common color space used in consumer cameras and devices, but other spaces could be used by operating with different matrices.

Then L_{w} has to be calibrated. Calibration is related to the linearity between HDR image and the absolute luminance, regardless the light conditions. This step is not biologically inspired but compensates for each camera individual properties: each camera is agnostic w.r.t. the scene acquired and will have their own encoding of the scenes which do not correspond necessary to the absolute luminance. Thus calibration is an important step since we do not know *a priori* how HDR image values differ from the real scene. In our case, it is also necessary since the photoreceptor adaptation model defined below needs absolute luminance values.

Some existing calibration methods relies on using meta-data as EXIF tags, GPS locations or light probes [37]. In this approach we use the heuristic proposed by [83, 80] for calibrating data, which does not need information about the environment where the image was taken upon. Following [83], the calibration process defines how to obtain an estimation of the real luminance $L(x, y, t)$ from $L_{\text{w}}(x, y, t)$:

$$L(x, y, t) = \frac{\alpha}{\bar{L}_{\text{w}}^{\log}(t)} L_{\text{w}}(x, y, t), \tag{1}$$

where the incoming data gets scaled by the ratio between the calibration coefficient α and the logarithmic average $\bar{L}_{\text{w}}^{\log}(\text{t})$ defined below.

This equation gives a way of setting the tonal range of the output image according to the so-called key of scene α. The key of the scene is a unit-less number which relates to whether the picture is subjectively

[1] Paper website: `https://team.inria.fr/biovision/tmobio`

[2] sRGB color space: https://www.w3.org/Graphics/Color/sRGB

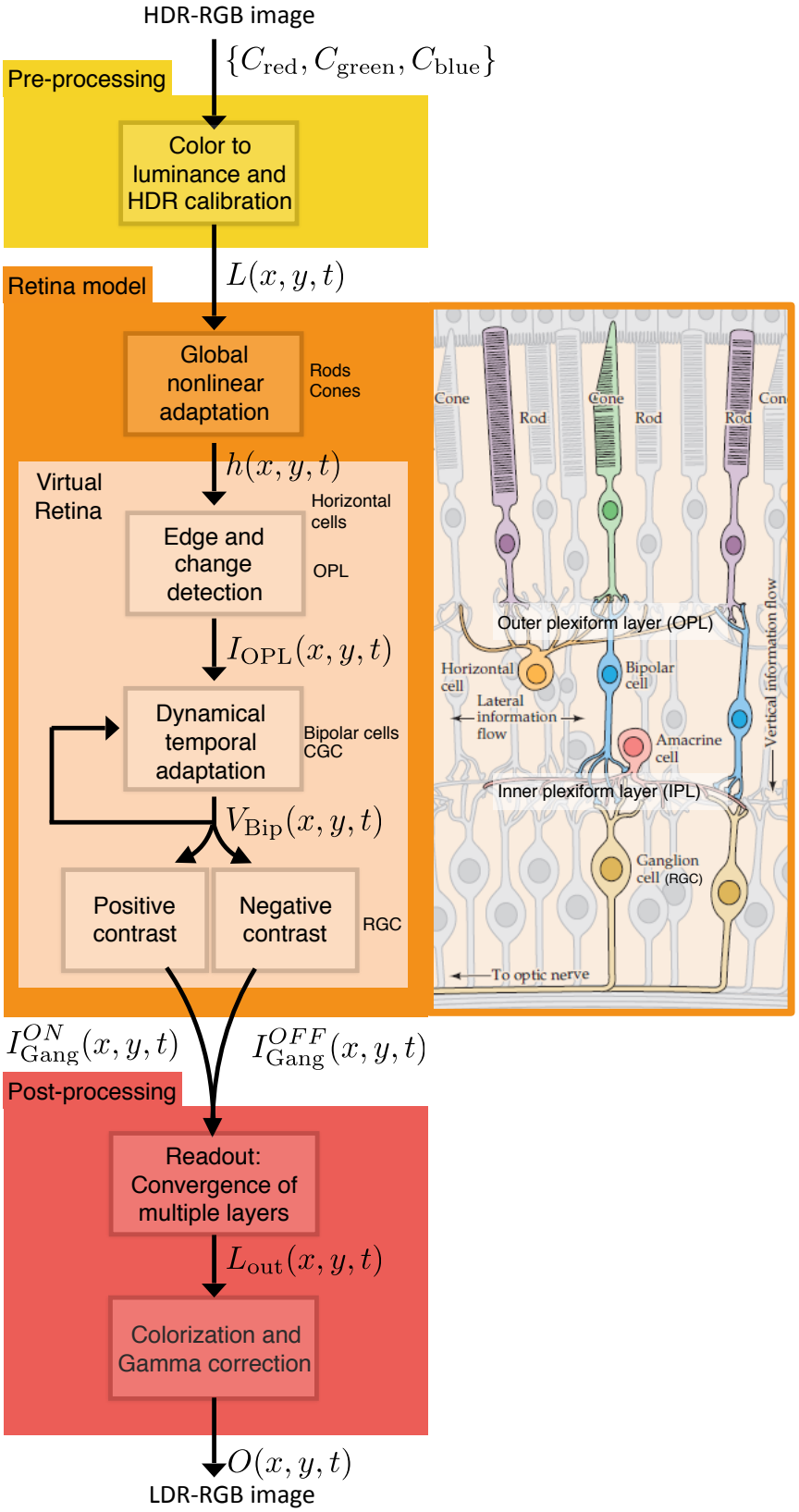


Figure 1: **Tone-mapping operator general pipeline**. Block diagram showing the three main stages of our method: Preprocessing, retina model extending VIRTUAL RETINA and post-processing. Retina main steps are illustrated on the right hand-side to remind that our retina models follows the main circuitry of primate retina (Courtesy [76])

dark, normal or bright. The underlying idea is that this ratio between α and $\bar{L}_{\mathrm{w}}^{\log}(t)$ will map the incoming luminance so that the calibrated output luminance would be correlated to the subjective notion of the brightness of the scene. The parameter α is generally user defined and set to a fixed value (0.18 for most of the cases), while the logarithmic average $\bar{L}_{\mathrm{w}}^{\log}(t)$ is an approximation of the key of the scene defined by

$$\bar{L}_{\mathrm{w}}^{\log}(t) = \exp\left(\frac{1}{N}\sum_{x,y\in\Omega}\log(\delta + L_{\mathrm{w}}(x,y,t))\right), \tag{2}$$

where N is the number of pixels in the Ω image space and δ is a small value in order to avoid numerical errors coming from zero-valued pixels.

Then equation (1) can be further improved by letting the key of the scene be automatically adjusted depending of the statistics of each image. Following [80], it can be obtained as follows:

$$\alpha(t) = 0.18 \cdot 4^{f(t)}, \tag{3}$$

with

$$f(t) = \frac{2\log_2(\bar{L}_w^{x,y,1\%}(t)) - \log_2(L_w^{\min,x,y,1\%}(t)) - \log_2(L_w^{\max,x,y,1\%}(t))}{\log_2(L_w^{\max,x,y,1\%}(t)) - \log_2(L_w^{\min,x,y,1\%}(t))}, \tag{4}$$

where $\bar{L}_w^{x,y,1\%}(t)$, $L_w^{\min,x,y,1\%}(t)$, and $L_w^{\max,x,y,1\%}(t)$ denote the average, minimum, and maximum values of the incoming data respectively over the spatial space, excluding 1% of the lightest and darkest pixels, statistically trimming values coming from overexposure in the camera or noise. This is very important since an image can have a distorted luminance distribution due to, e.g., overexposure. Note that in (3) the original value 0.18 is kept but is now modulated by $4^f(t)$, which takes into consideration the statistical features of the image through equation (4). The value 4 was empirically determined by the author [80].

2.3.2. Retina model

In this section, we try to present the different steps of the retina model in functional terms rather than in biological terms which would be less informative for our readers. For that purpose, we illustrate in Fig. 2 the effect of the retinal steps in TMO pipeline, and also provide an DRIM metric [2] for each stage in order to allow a better appreciation of the accumulative changes compared to the original HDR image. We refer the interested readers to the original paper describing VIRTUAL RETINA [101] to understand the biological foundation of the model. We also try to find analogies with previous works, not necessarily bio-inspired methods, to show that there are common ingredients found in artificial and biological models, even if their implementations differ (see also Sec. 2.4).

Photoreceptor adaptation: Global non-linear adaptation. Photoreceptors adjust their sensitivity over time and space according to the luminosity present in the visual scene, providing a first global adaptation for the incoming light [25]. This stage was not present in VIRTUAL RETINA original model which was not designed to process HDR images.

Here we consider the model by [25], where the authors proposed to fit empirical data coming from primate retinal cones. Their model describes photoreceptor activity as a function of the background and average luminance. Using their results, we showed how the luminance received by each photoreceptor can be mapped depending on the input luminance $L(x,y,t)$, following the nonlinear equation:

$$h(x,y,t;L) = \left(1 + \left(\frac{l_{1/2}(\bar{L}(t))}{L(x,y,t)}\right)^n\right)^{-1}, \tag{5}$$

where n is a constant which can be set according to experimental data to $n = 0.7 \pm 0.05$, $\bar{L}(t)$ is the instantaneous spatial average of the incoming luminance defined by $\bar{L}(t) = \int_{x,y} L(x,y,t)dxdy$, and $l_{1/2}$ is nonlinear function. Technical details to prove this relation (5) are given in the Appendix A (see also Fig. A.12 for an illustration of the function $l_{1/2}(.)$). Results about this stage are presented in Sec. 3.1.

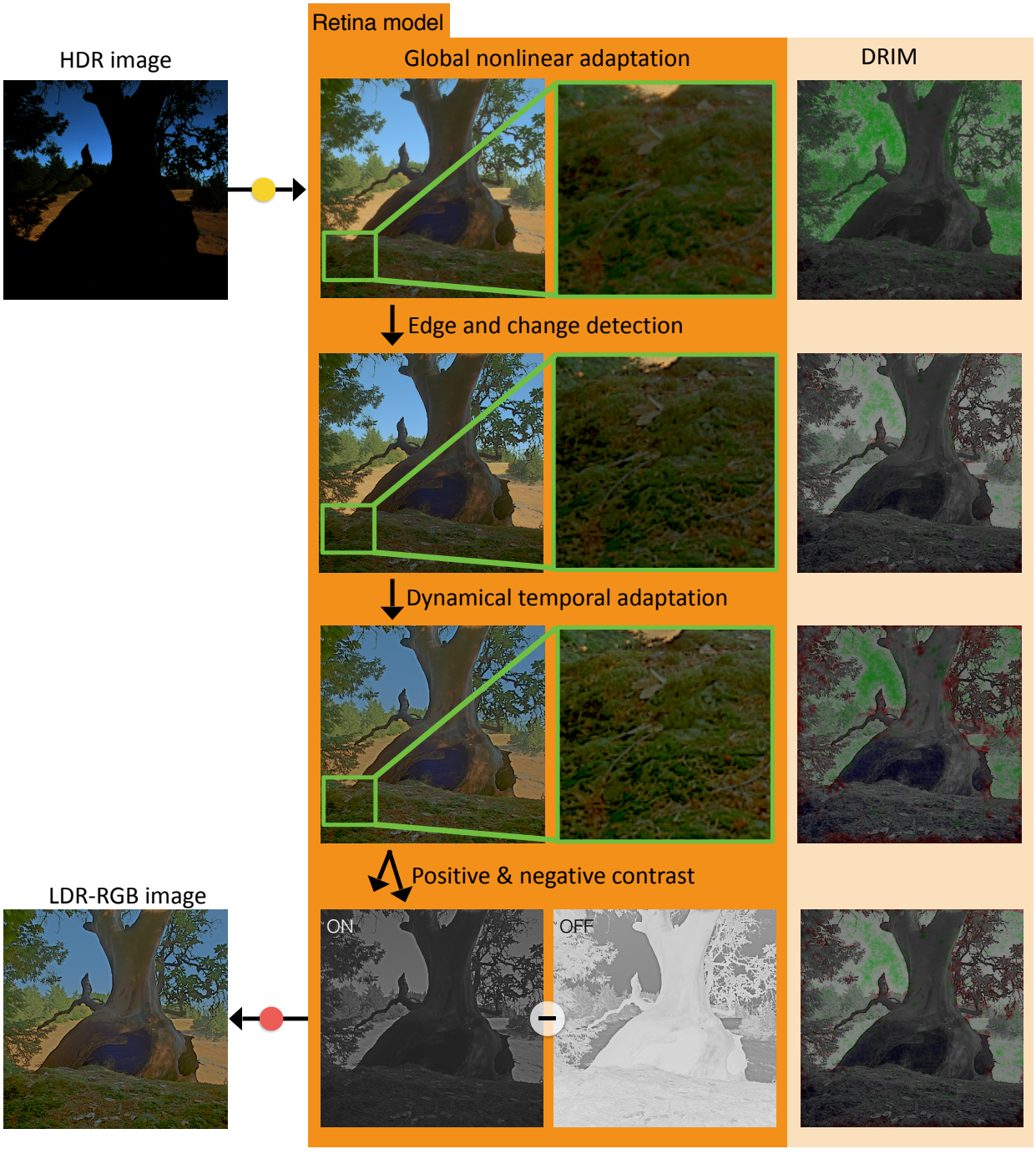


Figure 2: **Illustration of the "Retinal model" steps from TMO pipeline presented in Fig. 1 (pre- and post-processing are indicated by yellow and red colored circles respectively)**. Input image is *tree_mask* by *Image, Light and Magic*. On the first step the photoreceptor adaptation provides a compression of dynamic range. Then the OPL layer provides sharpening of the image. After that, the CGC layer improves the overall lightning reducing or amplifying the contrast locally. Finally, the readout is done by merging the two images describing positive and negative contrasts corresponding to ON and OFF ganglion cells. Note that for each step, except for contrast images, images have been colorized and gamma corrected to facilitate comparisons between different steps. These results were obtained using the model parameters presented on Table 1, except for $\lambda_A = 10000$ [Hz] and $\sigma_A = 1.2$ [deg]. For each step the respective DRIM [2] is presented, showing in green loss of contrast, blue amplification of contrast and red reversal of contrast against the input HDR image as reference. Remark that ganglion cell layer is not affecting the image contrast, but only the global brightness.

Our photoreceptor model can be related to the TMO proposed by [95] based on dynamic cone model. However, in our method dynamics is introduced later (namely at the CGC level) so that the photoreceptor adaptation behaves as a static nonlinear global operator.

Outer plexiform layer (OPL): Edge and change detection. This stage operates edge detection and temporal change detection given the light information already adapted by the photoreceptors ($h(x, y, t; L)$). In retinal circuitry, this is performed at the OPL level. OPL designates the first synapse layer in the retina making the connection between photoreceptors, horizontal cells and bipolar cells. Output of this stage is the I_{OPL} current. This band-pass behavior is obtained by center-surround interactions, which can be written:

$$I_{\mathrm{OPL}}(x, y, t) = \lambda_{\mathrm{OPL}}(C(x, y, t) - w_{\mathrm{OPL}} S(x, y, t)), \tag{6}$$

where $C(x, y, t)$ represents the center excitatory signal, calling for the photoreceptor adaptation and phototransduction and $S(x, y, t)$ is the surround inhibitory signal, which models the horizontal cell response. They are defined as follows:

$$C(x, y, t) = G_{\sigma_C} \overset{x,y}{*} T_{w_U,\tau_U} \overset{t}{*} E_{n_C,\tau_C} \overset{t}{*} h(x, y, t; L), \tag{7}$$

$$S(x, y, t) = G_{\sigma_S} \overset{x,y}{*} E_{\tau_S} \overset{t}{*} C(x, y, t), \tag{8}$$

where $*$ denotes convolution that can be either spatial ($\overset{x,y}{*}$) or temporal ($\overset{t}{*}$), G_σ is the standard Gaussian kernel, $T_{w,\tau}$ is a partially high-pass temporal filter, E_τ is an exponential filter to model low-pass temporal filtering and $E_{n,\tau}$ is an exponential cascade (Gamma filter)[3]. The balance between the center and surround contribution is given by w_{OPL} and it is normally close to 1. λ_{OPL} defines the overall gain of this stage. Parameters σ_C, n_C, τ_C, w_U, τ_U, σ_S and τ_S are other constant parameters.

Equation (6) with (7)–(8) results in a spatiotemporal non-separable filter. It extends the only-spatial classical model of the retina as a difference of Gaussians (DoG). As such, the I_{OPL} is a real valued function corresponding to the presence of contrasts in the scene either a spatial contrast (with sign indicating if there is a bright spot other a dark background or the opposite) or a change in time (with sign also indicating the nature of the temporal transition). This real-valued function will lead to the two rectified components defining ON and OFF cells in the ganglion cell layer (see below). This is a classical model in neuroscience so that all its parameters are well understood and can be set to default values given by biological constrains. Results about this stage are presented in Sec. 3.2.

Note that this DoG-like operation done in the OPL layer in conjunction with the dynamic range compression performed by the photoreceptor adaptation generate an operator which can be qualitatively compared to the bilateral filter TMO proposed by [26] in the sense that there is an implicit estimation of edges.

Contrast Gain Control layer (CGC): Dynamical temporal adaptation. Contrast gain control, or contrast adaptation, is the usual term to describe the influence of the local contrast of the scene on the transfer properties of the retina [90, 97, 45, 85, 4]. In VIRTUAL RETINA, an original model was proposed for fast contrast gain control, an intrinsically dynamic feature of the system, which has a strong and constant influence on the shape of retinal responses, and very likely on our percepts. It is an effect intrinsically nonlinear, and dynamical. Dynamically, it modulates the gain of the transmission depending on the local contrast.

In VIRTUAL RETINA, given the input current $I_{\mathrm{OPL}}(x, y, t)$, the membrane potential of bipolar cells denoted by V_{Bip} is evolving according to:

$$\frac{dV_{\mathrm{Bip}}}{dt}(x, y, t) = I_{\mathrm{OPL}}(x, y, t) - g_A(x, y, t) V_{\mathrm{Bip}}(x, y, t), \tag{9}$$

with

$$g_A(x, y, t) = \left(G_{\sigma_A} \overset{x,y}{*} E_{\tau_A} \overset{t}{*} Q(V_{\mathrm{Bip}}) \right)(x, y, t) \text{ and } \; Q(\nu) = g_A^0 + \lambda_A \nu^2,$$

[3]Expressions are: $E_\tau = \exp(-t/\tau)/\tau$ and $E_{n,\tau}(t) = (nt)^n \exp(-nt/\tau)/\left((n-1)!\tau^{n+1}\right)$

where σ_A, τ_A, g_A^0 and λ_A are constant parameters. This model has been rigorously evaluated by comparisons with real cell recordings using specific stimuli.

So, in practice this stage translates in contrast equalization, which we see as an image enhancing capability and a temporal coherence source for videos. Figure 2 illustrates this effect as the increase of brightness in some areas. As in the former stage, all its parameters can be set to default values to fit biological constrains. Results about this stage are presented in Sec. 3.3.

Ganglion cells layer. In the last layer of the retina, namely ganglion cells layer, two main processing stages happen. The first consists in transforming the potential of bipolar cells into an electric current. The second consists in transforming this current into spikes which is how information is encoded in the end by the retina to be transmitted to the visual cortex through the optic nerve. In this paper we will only consider the first stage since the spike-based encoding will not be exploited.

From a more functional perspective, it is in this last stage that the response of ganglion cells can be interpreted in terms of features in the observed scene. There are about 20 ganglion cell types reporting to the brain different features [59]. In this paper we focus on two major classical types, namely ON and OFF cells which respond to positive and negative spatiotemporal contrasts. As explained above, the emergence of contrast sensitivity comes from the Outer Plexiform Layer where center-surround processing occur (see Eq. (6)).

Following the model proposed in VIRTUAL RETINA, ON and OFF ganglion cell currents (resp. denoted by $I_{\text{Gang}}^{\text{ON}}(x,y,t)$ and $I_{\text{Gang}}^{\text{OFF}}(x,y,t)$) are obtained by applying a nonlinear rectification to the bipolar potential to either keep the positive or negative parts. This can be written as follows:

$$I_{\text{Gang}}^{\text{ON/OFF}}(x,y,t) = \mathcal{N}\left(\varepsilon\ V_{\text{Bip}}(x,y,t)\right), \tag{10}$$

with

$$\mathcal{N}(V) = \begin{cases} \frac{i_G^0}{1-\lambda_G(V-v_G^0)/i_G^0} & \text{if } V < v_G^0, \\ i_G^0 + \lambda_G(V - v_G^0) & \text{if } V > v_G^0, \end{cases} \tag{11}$$

where ε defines the output polarity ($\varepsilon = 1$ for ON cells and $\varepsilon = -1$ for OFF cells), i.e., changing the ε parameter in (10), one defines the two ganglion cells currents $I_{\text{Gang}}^{\text{ON}}(x,y,t)$ and $I_{\text{Gang}}^{\text{OFF}}(x,y,t)$, describing respectively positive and negative contrasts. The function $\mathcal{N}(V)$ is a non-linear function which rectifies $I_{\text{Gang}}^{\text{ON/OFF}}(x,y,t)$ using a smooth curve. Such rectification is a very common feature in neural modeling and in retinal models. It reflects static nonlinearities observed experimentally in the retina. The parameter v_G^0 is the "linearity threshold" between the two operation modes while λ_G is the slope of the rectification in the linear part. Note that (10) is a simplification of the original model in which additional convolutions were added to fit biological data.[4]

Results about this stage are presented in Sec. 3.4 where we show the influence of the parameter λ_G.

Remark 1. ***Are ON an OFF circuits symmetric?*** *In the current model and following [101], ON and OFF circuits are symmetric in the sense that they were obtained by simply changing the polarity ε. However, this is a simplified assumption with respect to real retina properties. Indeed, there are physiological and mechanistic asymmetries in retinal ON and OFF circuits [17, 106, 5, 47]. For example, across species, it has been observed that OFF cells (compared to ON cells) are more numerous, sample at higher resolution with denser dendritic arbor and respond to contrast with strong rectification. Differences between the two circuits have been also observed depending on the overall light condition, e.g., photopic or scotopic conditions [72, 68, 92]. Interestingly, this asymmetry could arise from the asymmetries between dark versus bright regions in natural images [5, 79]. For example, in [79] the authors considered the distribution of local contrast in 100 natural images and suggest that because of the skew observed in that distribution (natural scenes contain more negative than positive contrast; see also ON and OFF ganglion cells response given in Fig. 2), the*

[4] In the original model, there were additional spatial and temporal convolutions (see Eq. (14) from [101]). These terms were useful to fit biological data but in our context they were introducing some blur so that we decided to discard them.

system should devote more contrast levels to encode the negative contrasts for example (e.g., OFF channels should be more rectified). All these results suggest that an interesting avenue for further study is to model explicitly these asymmetries to better account for natural image statistics (see discussion in Sec. 4). This is why we introduced the ON and OFF circuits here although there role in the present study remains limited (slight impact on global luminance) and we refer to Sec. 3.4 where we show a preliminary result where some asymmetry is introduced between ON and OFF circuits.

Readout population activity. Given ON and OFF responses, one needs to define how to readout an activity based on different population responses. Here we chose a simple way to combine the activity generated by both the ON and OFF circuits:

$$L_{\text{out}}(x, y, t) = I^{\text{ON}}_{\text{Gang}}(x, y, t) - I^{\text{OFF}}_{\text{Gang}}(x, y, t), \tag{12}$$

where $L_{\text{out}}(x, y, t)$ represents the output radiance map which can now be colorized, normalized, quantized and gamma corrected for the output LDR file, as described in final step below.

Colorization and Gamma correction. Up to now, our method operates on the luminance map of the input image. Once the luminance map has been corrected, it is necessary to colorize it. To do this we apply the method proposed by [57] which involves linear interpolation between chromatic input information ($\{L_i(x, y, t)\}_{i=\text{red, green, blue}}$), the respective achromatic input information ($L(x, y, t)$) and the achromatic output information $L_{\text{out}}(x, y, t)$). To obtain the output in linear space $O_{\text{lin},i}$, each chromatic channel is colorized as follows:

$$O_{\text{lin},i}(x, y, t) = \left(\left(\frac{L_i(x, y, t)}{L(x, y, t)} - 1 \right) s + 1 \right) L_{\text{out}}(x, y, t), \tag{13}$$

where s is the color saturation factor, in our implementation set by default to 1.

The model also considers a gamma correction step. Luminance values coming from a HDR image are usually in a linear scale, which does not map well with the color spaces of computer monitors mostly having non linear spaces. Moreover these response curves are usually described by an power law, so we need to apply the inverse operation to keep the desired appearance correctly on screen. For each channel chromatic channel, a standard gamma correction is applied:

$$O_i(x, y, t) = \begin{cases} 0 & 255 \cdot O_{\text{lin},i}^{1/\gamma} \leq 0, \\ \lfloor 255 \cdot O_{\text{lin},i}^{1/\gamma} + 0.5 \rfloor & 0 < 255 \cdot O_{\text{lin},i}^{1/\gamma} \leq 255, \\ 255 & \text{otherwise}, \end{cases} \tag{14}$$

where the choice of the gamma factor normally depends on the screen, (default: $\gamma = 2.2$). The factor of 255 present in the equation is related to the maximum possible value that can be stored in a common LDR image format like PNG or JPEG with a resolution of 8 bits per pixel.

2.4. Relation with previous work

Given the nature of our model, in this section we first discuss a selection of models inspired by visual system properties and architecture. We found a range of models from psychophysical inspiration (e.g., based on fitting psychophysical curves for light adaptation) to more bio-plausible models of the visual system (e.g., mathematical model of the retinal circuitry).

One of the earliest examples, [78] developed an TMO based on the retinex algorithm. This yields a local non-linear operator that relies on the work of [50], who suggested that both the eye and the brain are involved in color processing, particularly in the color constancy effect. The TMO proposed by [78] process each color channel independently and finally merged using a color restoration method proposed by the author. Their method has two modes of operation: single and multi-scale. At single scale each color channel is convolved with Gaussian kernels of different sizes. Each pixel (center) is combined with the blurred image (surround) in a nonlinear manner defining thus different center/surround interactions at different spatial

scales. At multi-scale the resulting center/surround interactions are weighted and summed obtaining the equalized output image. The selection of the weights is based on a heuristic considering either small or large surrounds.

In [81], the authors developed a TMO inspired in the adaptation mechanism present in photoreceptors. This TMO describes the automatic adjustment to the general level of brightness that color-photoreceptors (cones) have. The core of their algorithm is the adaptation of a photoreceptor model, which is mainly implemented as a variation from the Naka-Rushton equation[5]. Their model lies on a global adaptation of the image fitting the Naka-Rushton parameters in a global (from average scene luminance) or local (from color average luminance) manner. This method does not allow luminance coherence over time because of its static formulation (each frame is processed independently). In our case, photoreceptor adaptation is done frame by frame, and luminance coherence is obtained by the temporal filtering in the CGC layer.

A front-end retina-like model for computer vision tasks was proposed by [9, 39] and applied for tone mapping images and video [8]. They modeled the retina as two adaptation layers. The first layer is related to photoreceptor adaptation and the second one to ganglion cells. Between the two main layers lies a non-separable spatio-temporal filter modeling the synaptic Outer Plexiform Layer (OPL), which summarizes the interaction between photoreceptors and horizontal cells of the retina. For the photoreceptor adaptation the authors propose a non-linear local process in which each photoreceptor response depends on the incoming luminance inside a local neighborhood. The OPL layer is then used as spectral whitening and contrast enhancement. In a similar manner, the ganglion cell layer reuses the nonlinear mechanism used for the photoreceptor adaptation with different parameters. Specifically, the authors consider a smaller neighborhood around each pixel to compute the average luminosity. Color processing is done through a multiplexing/demultiplexing scheme based on the Bayer color sampling, normally found in cameras. Regarding HDR video, the authors use the temporal aspect of the OPL spatio-temporal filter to provide temporal coherence through the sequence. This model shares some features with Virtual Retina model [101], for instance, the modeling of the OPL as a non-separable spatio-temporal filter. [8] directly connects the OPL output to the nonlinearity proposed in the ganglion cell layer, while in Virtual Retina OPL output gets into a Contrast-Gain Control stage before the ganglion cell layer.

Finally, even though is not properly a TMO, in [107], the authors proposed a method for de-hazing static images that relies on a thorough retina model, from photoreceptors to retinal ganglion cells. In this model, photoreceptors, bipolar, amacrine and ganglion cells are modeled as a cascade of Gaussian and difference of Gaussian filters through several signal pathways, representing the intricate relationships between colors and global versus local adaptation. We mention this work here because our model shares the modeling of those biological layers, except for amacrine cells. Also, their model does not account for any dynamical Contrast Gain Control which is an important processing stage in our method.

Compared to general video tone mapping approaches, it is worth mentioning that our method does not need optical flow estimation, as in [3, 38, 55] where the authors use it stabilize scenes with small motion, with the limitation of producing artifacts in scenes with larger displacements [43]. Other video TMOs apply transforms that change over time but applied uniformly in space [32, 74, 42, 95, 56, 82]. They perform well in terms of temporal coherence but can produce a loss in detail, whereas our model take into account local contrast information in space, which better preserves spatial details.

3. Results and comparisons

In this section we evaluate and compare our bio-inspired synergistic TMO on both static images and videos. Remark that if the input is a static image, it is considered as a video by simply repeating that image other time.

Unless specified otherwise, all model parameters used to process HDR images are set as given in Table 1. It is important to emphasize that a majority of these parameters have been set thanks to biological

[5]Sigmoid function which computes the response of the photoreceptor based on the incoming brightness and the semi-saturation constant

Table 1: Default parameters used in our implementation. Note that all parameters and variables presented in the VIRTUAL RETINA model are expressed in reduced units and not physical units (see justification in Sec. 2.1.2 from [101]). As a consequence, for example, the dimension of currents is expressed in Hertz and the overall gain of the center-surround filter λ_{OPL} is expressed in Hertz per unit of luminance.

Parameter	Value	Parameter	Value	Parameter	Value
n	0.5	$i_{1/2}$	52000 $\left[\text{P s}^{-1}\right]$	σ_C	0.03 [deg]
τ_C	0.01 [s]	n_C	2	τ_U	0.1 [s]
w_U	0.8	σ_S	0.2 [deg]	τ_S	0.01 [s]
λ_{OPL}	10 $\left[\text{Hz lum}^{-1}\right]$	w_{OPL}	0.55	dt	0.005 [s]
t_f	0.2 [s]	λ_A	100 [Hz]	σ_A	0.2 [deg]
τ_A	0.0005 [s]	g_A^0	5 [Hz]	v_G^0	0
i_G^0	80 [Hz]	λ_G	100 [Hz]	s	1
γ	2.2				

constraints. Note that since all gaussian standard deviation parameters were defined in visual angle degrees in VIRTUAL RETINA, we used a conversion of five pixels per visual angle degree, or [p.p.d.]. This value was kept constant for different image sizes. Concerning temporal constants, we assume that videos are at 30fps and between each frames we do six iterations of the CGC equation (9). For a static input image, we let the iterative process converge towards a final image (in practice, about five repetitions of the frame, i.e., 30 iterations). We used a time step of $dt = 0.5$ [ms].

Sections 3.1 to 3.5 deal with static images. We will show the influence of some parameters in each step to show their impact and benchmark our method with respect to the state-of-the-art. Section 3.6 gives an example of result on a real video sequence. We show how the temporal convolutions provide temporal coherence in the output.

3.1. Photoreceptor adaptation: Global non-linear adaptation

Figure 3 shows the effect of photoreceptor adaptation. In Fig.3(a) we show how an input radiance map $L(x, y, .)$ (converted to an LDR image using a linear mapping) is transformed into $h(x, y, .; L)$ image defined in (5). This is done by a global non-linear operator effectively compressing the whole incoming data range to the interval $]0, 1[$, step necessary for the VIRTUAL RETINA model to work properly. The behavior of this nonlinearity amplifies the cases where the luminance is higher compared to the mean spatial luminance of the scene, nonlinearly computed through $l_{1/2}(\cdot)$, i.e., overexposed regions. In Fig. 3(b) we show the absolute difference between $h(\cdot)$ and $L(\cdot)$[6]. Note that the main differences between the two images are located in the overexposed regions present in the stained glass windows.

A comparison between different values of the exponential parameter n in Eq. (5) is shown in Fig. 4. In the experimental fit of [25] this constant was set to $n = 0.7 \pm 0.05$. One can change n to increase or decrease the effect of photoreceptor adaptation. From our tests, we found that $n = 0.5$ gives consistently good results.

3.2. Outer plexiform layer (OPL): Edge and change detection

In Fig. 5 we analyze the impact of the parameters from Eq. (6) defining the center-surround interaction. First we vary the relative weight value of w_{OPL} (experimentally estimated near 1 for mammal retinas, and in general in the range of $[0, 1]$). High values of w_{OPL}, e.g., 0.9 excessively enhances borders specially when the size of the surround, given by the σ_S parameter, increases. So, we have experimentally set the parameter w_{OPL} with a default value of 0.55, which performs consistently well in our tests. Then the size of the surround produces a high impact on the resulting I_{OPL} layer. In Fig. 5 we show that increasing the value of σ_s induces artifacts for high values of w_{OPL} specially in borders with a high difference of luminosity, mainly because saturation in the $I_{\text{OPL}}(x, y)$ value. The effect of σ_S additionally enhances details on regions with

[6]Both images were normalized between [0,1] to compute the absolute difference between them

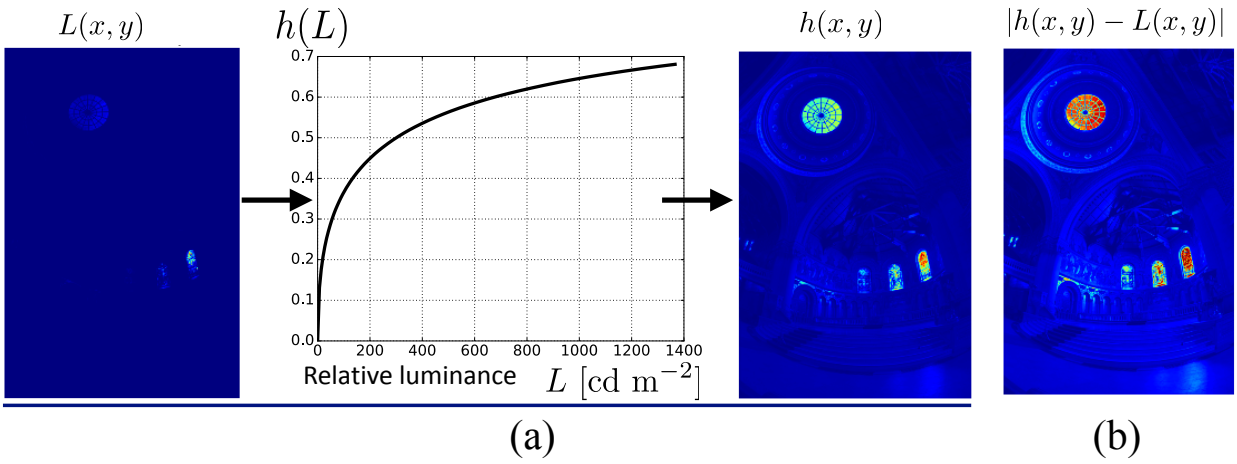


Figure 3: **Effect of global non-linear photoreceptor adaptation** **(a)** Effect of the non-linear operator $h(\cdot)$ defined in (5) over the input luminance image L. The image obtained after the photoreceptor adaptation is represented in $h(\cdot)$. **(b)** Difference of the normalized signals $|h - L|$.

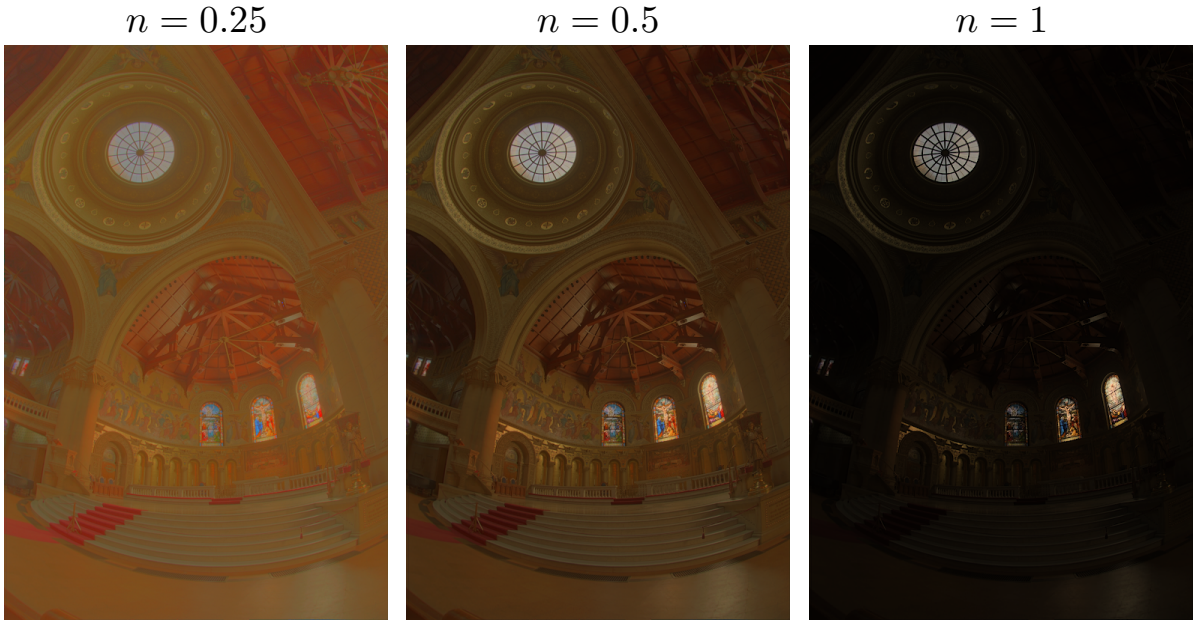


Figure 4: **Effect of the parameter n in the photoreceptor adaptation** (5). This parameter has an effect in terms of compression of the incoming luminance map. When n is low, compression is exaggerated resulting in a arguably distorted image. When n is close to one, function $h(x, y, t; L)$ behaves as a quasi-linear operator.

low contrast when there is no saturation (low values of w_{OPL}). The default value in our implementation is $\sigma_S = 0.1$ [deg], approximately the modeled value for primate retinas. This value shows a good compromise between edge enhancement and image distortion, while being biologically accurate.

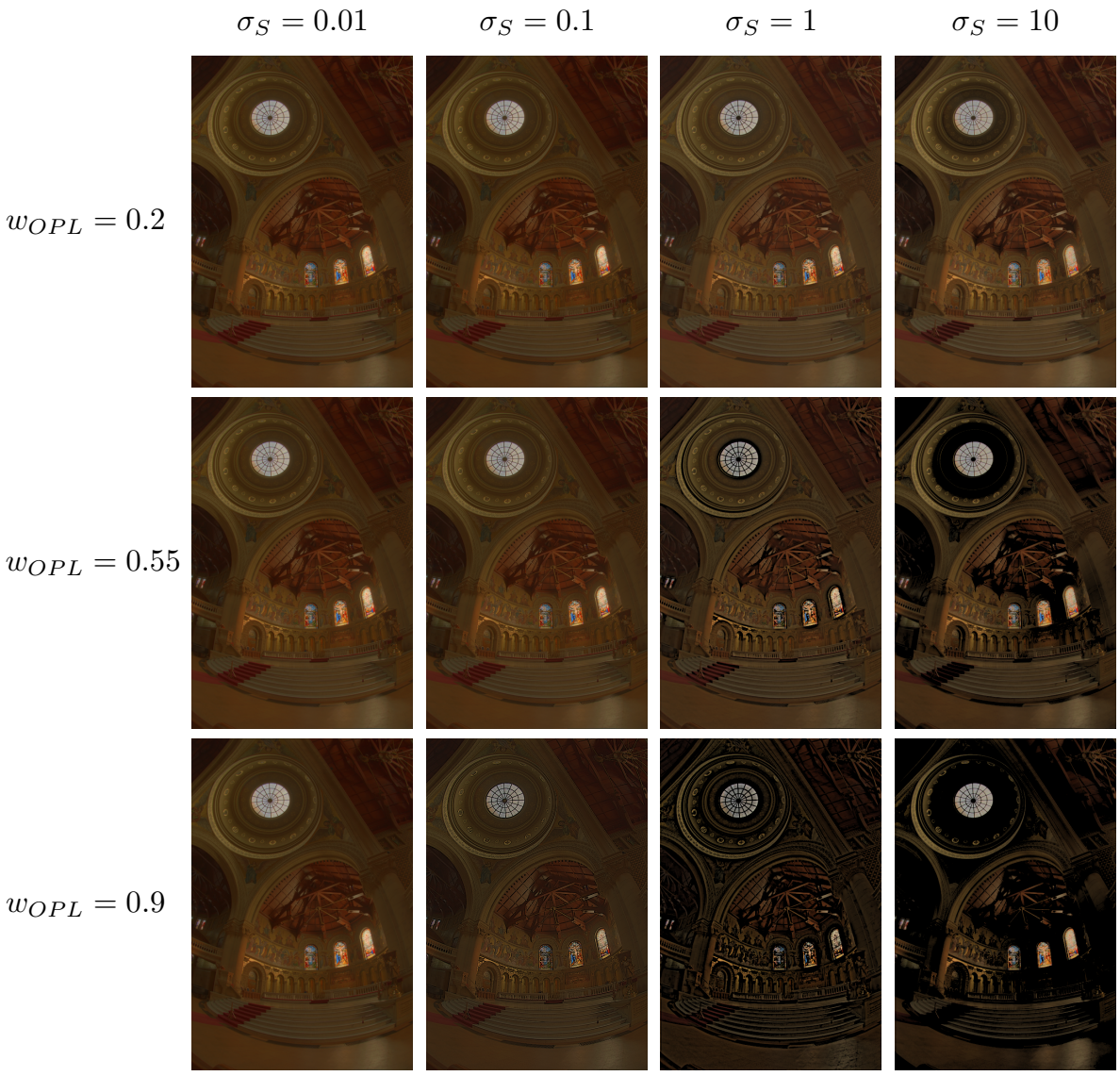


Figure 5: **Resulting $I_{\mathbf{IOPL}}$ for different values of σ_S and $w_{\mathbf{OPL}}$** (6). The relative weight w_{OPL} highlights edges in the input image according to the local average described by the surround signal $S(x, y)$. This becomes more prominent with a large values of σ_S.

3.3. Contrast Gain Control layer (CGC): Dynamical temporal adaptation

In Fig. 6 we show the effect of Contrast Gain Control (CGC) varying parameters. CGC layer enhances dynamically local brightness and local contrast in the image. According to (9), there are two parameters that directly impact the value of $V_{\mathrm{Bip}}(x, y, t)$: λ_A, the overall gain of the feedback loop and σ_A, the standard deviation of the spatial gaussian blur. In Fig. 6 we focus on these two parameters. Similarly to what was reported by [101], the gain in the feedback loop given by λ_A translates into more intense differences with

respect to the I_{OPL} input signal, obtaining as a result brighter images. The value of $\lambda_A = 1$ [Hz] yields little differences regardless of the choice of σ_A. The effect of σ_A is evident when the value of λ_A increases, and in this case the gaussian blur produced by σ_A generates local enhancement specially observed in overexposed regions such as the stained glass windows. After this analysis we decided to set the default values of these parameters to $\lambda_A = 10^4$ [Hz] and $\sigma_A = 1.2$ [Hz].

In Fig. 7 we illustrate how the dynamics of CGC layer work. We denote by $V_{\mathrm{Bip}}^{n_i}$ the value of the map V_{Bip} after n_i iterations of the discretized differential equation (9). We show how the CGC layer progressively enhances details in time. In Fig. 7(a) we show the image of $V_{\mathrm{Bip}}^{n_i}$ for $n_i = 0$, 10 and 30 iterations, which can be considered as the steady-state. Convergence of the differential equation to the steady state is illustrated in Fig. 7(c) showing the relative difference between two successive iterations. Stronger variations are observed during the first 10 iterations. Numerically, we chose $t_f = dt \cdot 20 = 0.2$ [s], as the default stop time for CGC dynamics over an static image. The effect of the CGC is more noticeable for the windows of the altar where most of the light of the scene comes from (see zoom). This is also visible from Fig. 7(b) where we show the difference between $V_{\mathrm{Bip}}^{n_i}$ at initial state ($n_i = 0$) and at convergence (here $n_i = 0$). CGC compensates the contrast locally, thus allowing for details to be enhanced, as the edges of the stained glass. Also, the additional compression in dynamic range translates into an overall brighter image, shown on the walls around the windows.

3.4. Ganglion cells layer

In Fig. 8 we study the influence of parameter λ_G in the ganglion cells layer (12). This parameter represents the slope of the rectification in the linear part.

In Fig. 8**(a)** we show results for different values of λ_G. Note that in order to enhance the effect of this stage, we pushed the CGC parameters to be $\sigma_A = 0.2$ [deg] and $\lambda_A = 10^2$ [Hz]. High values of λ_G (10^6 [Hz]) distort the output image with an excessive brightness, while low values of $\lambda_G (\in [10^0, 10^4]$ [Hz] approx.) improve the overall brightness without overexposing the scene. Following our motivation to keep the TMO biologically plausible, the value chosen as default for this gain is $\lambda_G = 100$ (see [102], Midget cells).

In Fig. 8**(b)** we propose a preliminary result to illustrate the effect of introducing asymmetries between ON and OFF circuits (see Remark 1 in the paragraph Ganglion cell of Sec. 2.3.2). This is done by choosing different values of slope of the sigmoid function (i.e., λ_G) to obtain the ON and OFF responses. Namely, we choose a fixed value to estimate the ON part ($\lambda_G^{\mathrm{ON}} = 100$) while we vary the value for the OFF part (λ_G^{OFF}). Increasing λ_G^{OFF} has a directly proportional effect in the compression rate over the channel, thus diminishing the contribution to the output luminance map. This effect can also be understood as the effective reduction of global contrast in the image.

3.5. Comparison with other methods

In Fig. 9 we visually compared our results to linear compression, logarithmic compression, Retinex model [78], histogram adjustment [51], bilateral filter-based approach [26], photo-receptor model [81] and bio-inspired spatio-temporal model [8]. We used implementations provided in [84]. These were run under a Linux machine using our C++ implementation. The TMO proposed in this paper was implemented using a default set of parameters, which can be found in Table 1. We used four HDR images in total, all of which are already properly calibrated. Every tone mapped image was obtained without intervention of the default parameters proposed in the respective models so that a fair comparison can be done. Results show that our method gives promising results in general for the four images selected. It manages to compress the luminance range without overexposure. For *Bristol Bridge*, a typical twilight outdoors scene, our method allows to see both the woods and the sky. In *Clock Building*, both the edifice details and the outdoors lighting are preserved. In *Memorial* the inside of the altar is recovered while conserving the details from the stained glass. Finally in *Waffle House* both the indoors of the building and the car can be seen without too much specular lighting noise, which can be seen over the roof of the building.

Note that we carefully verified that methods were tested in the same conditions regarding input calibration and gamma correction to obtain the final output. This is important in order to provide a fair comparison ground for every TMO. Concerning the calibration step (see Sec. 2.3.1) this step is not a priori

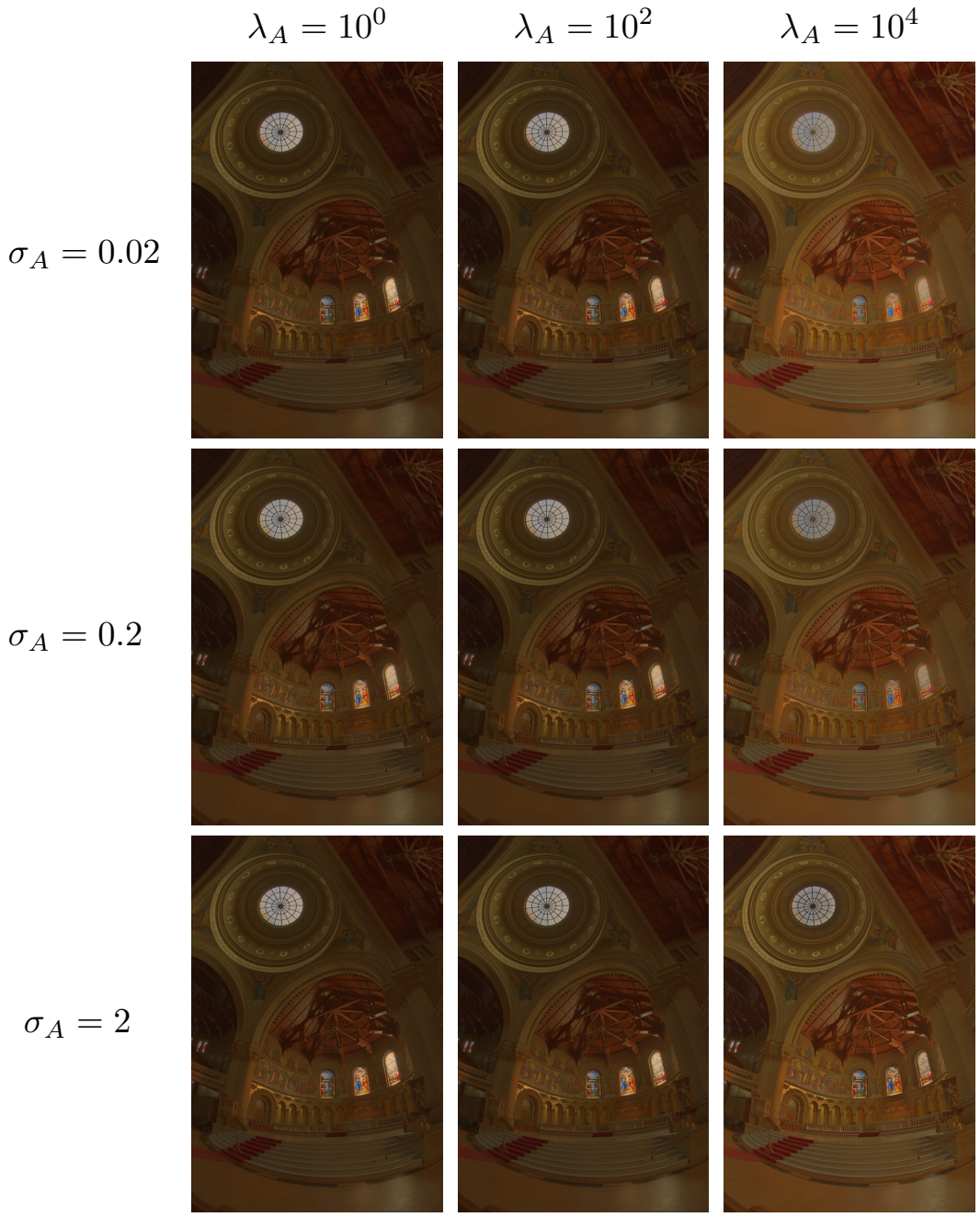


Figure 6: **Effect of CGC parameters on $V_{\mathbf{Bip}}$**. This figure summarizes the effect of λ_A and σ_A parameters on the Contrast Gain Control layer behavior. The gain of the loop, given by λ_A, translates into an improved overall brightness in the image, while an increase of the standard deviation σ_A value enhances details. The parameter values chosen as default in this model are shown in Table 1.

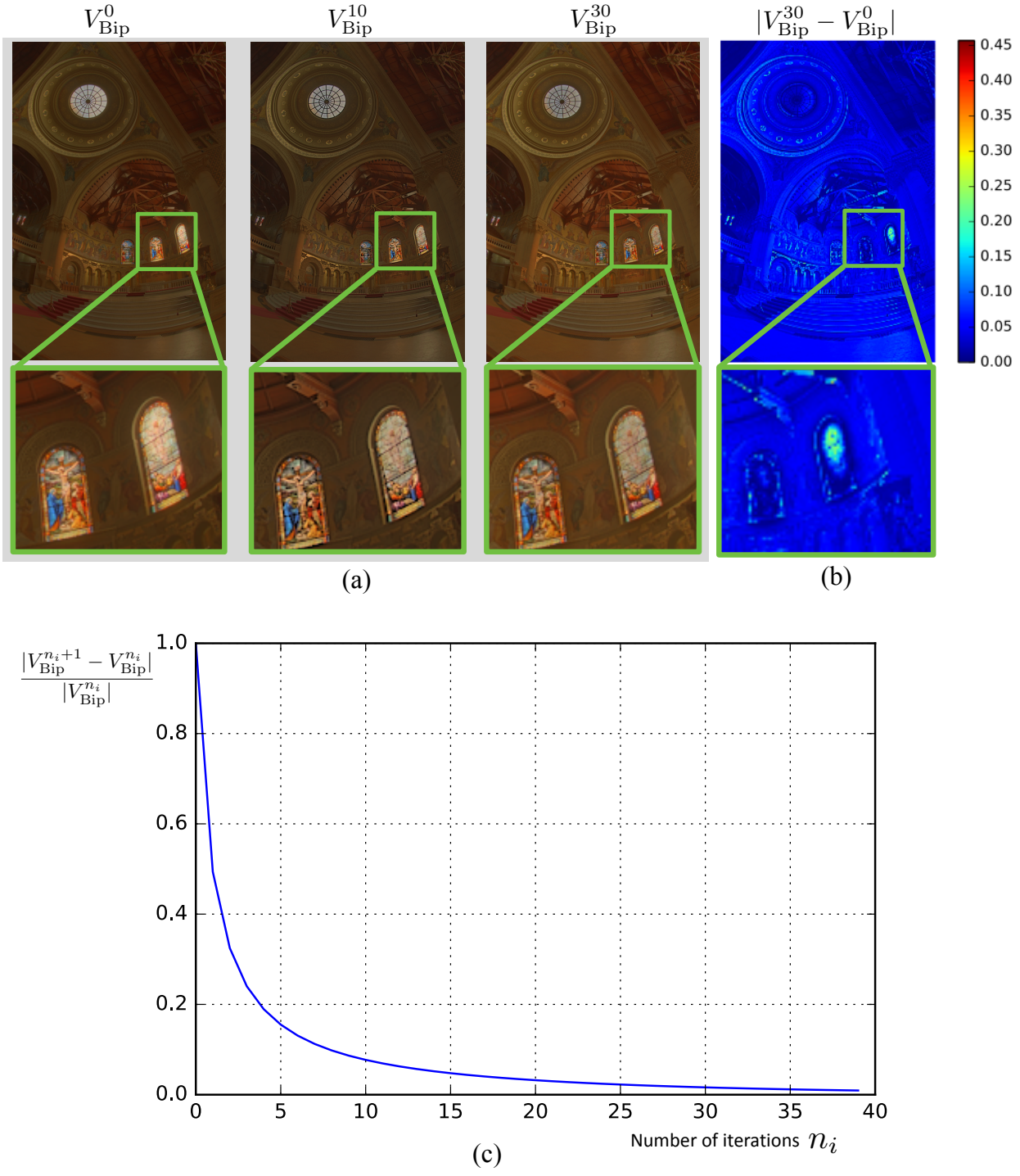


Figure 7: **Effect of CGC in local contrast enhancement**. Comparison between the input and output of the CGC stage for different iterations number. **(a)** Evolution of $V^{n_i}_{\text{Bip}}$ after $n_i = \{0, 10, 30\}$ iterations (note that resulting image was colorized and gamma corrected for display). **(b)** A colorized map of the differences between the iteration 30 and the initial state at iteration 0. **(c)** The mean squared error computed between two consecutive iterations of $V^{n_i}_{\text{Bip}}$. After approximately 15 iterations the changes on the scene are not noticeable.

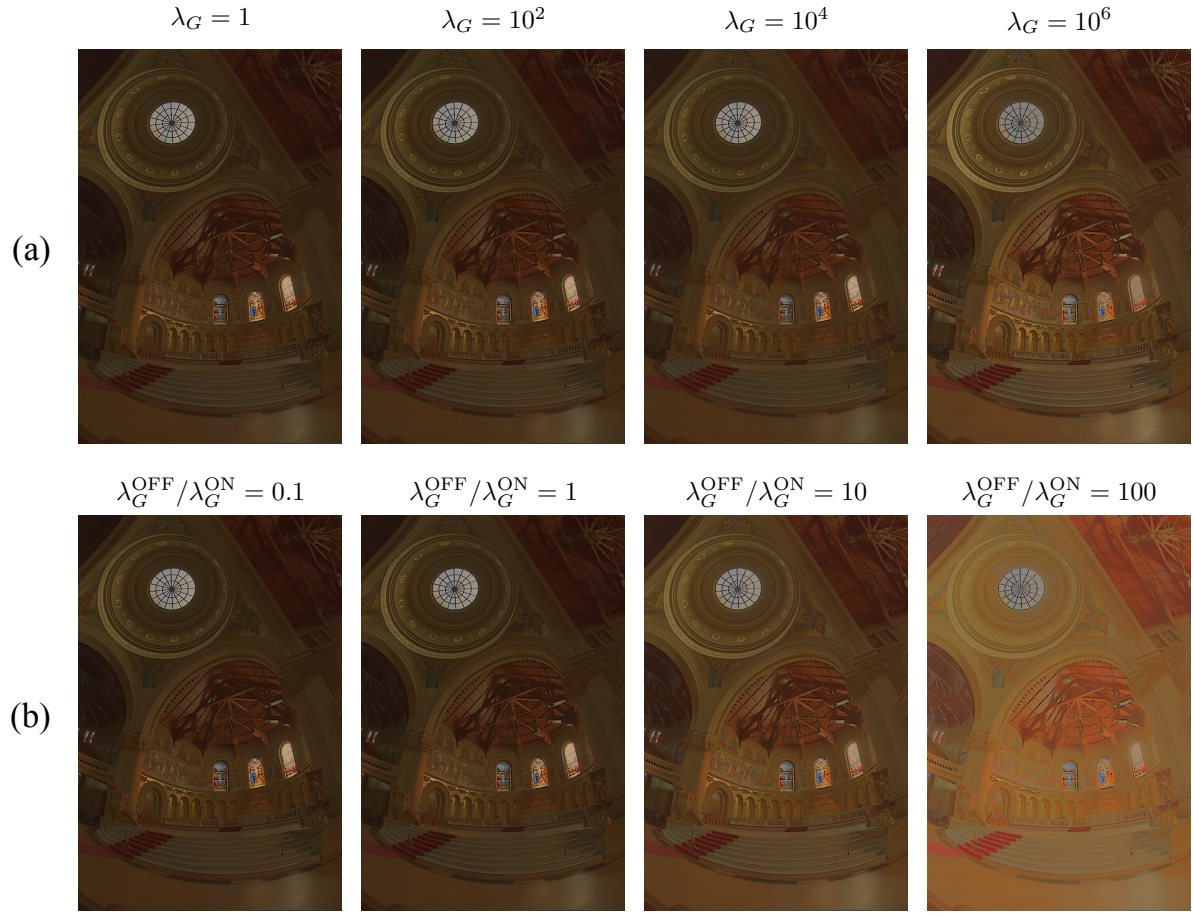


Figure 8: **Effect of λ_G on the output of the Ganglion cells layer $L_{\mathbf{out}}(x, y)$. (a)** Testing different values of λ_G. High values of λ_G ($> 10^6$ [Hz]) saturate image brightness: see distortions in the ceiling and the stained glass windows. Lower values of λ_G, on the contrary, improve the global perception of the scene. **(b)** Introducing asymmetry between ON and OFF circuits. Different values of λ_G are used for each circuit: λ_G^{ON} is fixed; λ_G^{OFF} is changed. For bigger values of λ_G^{OFF} there is a clear loss of global contrast across the image.

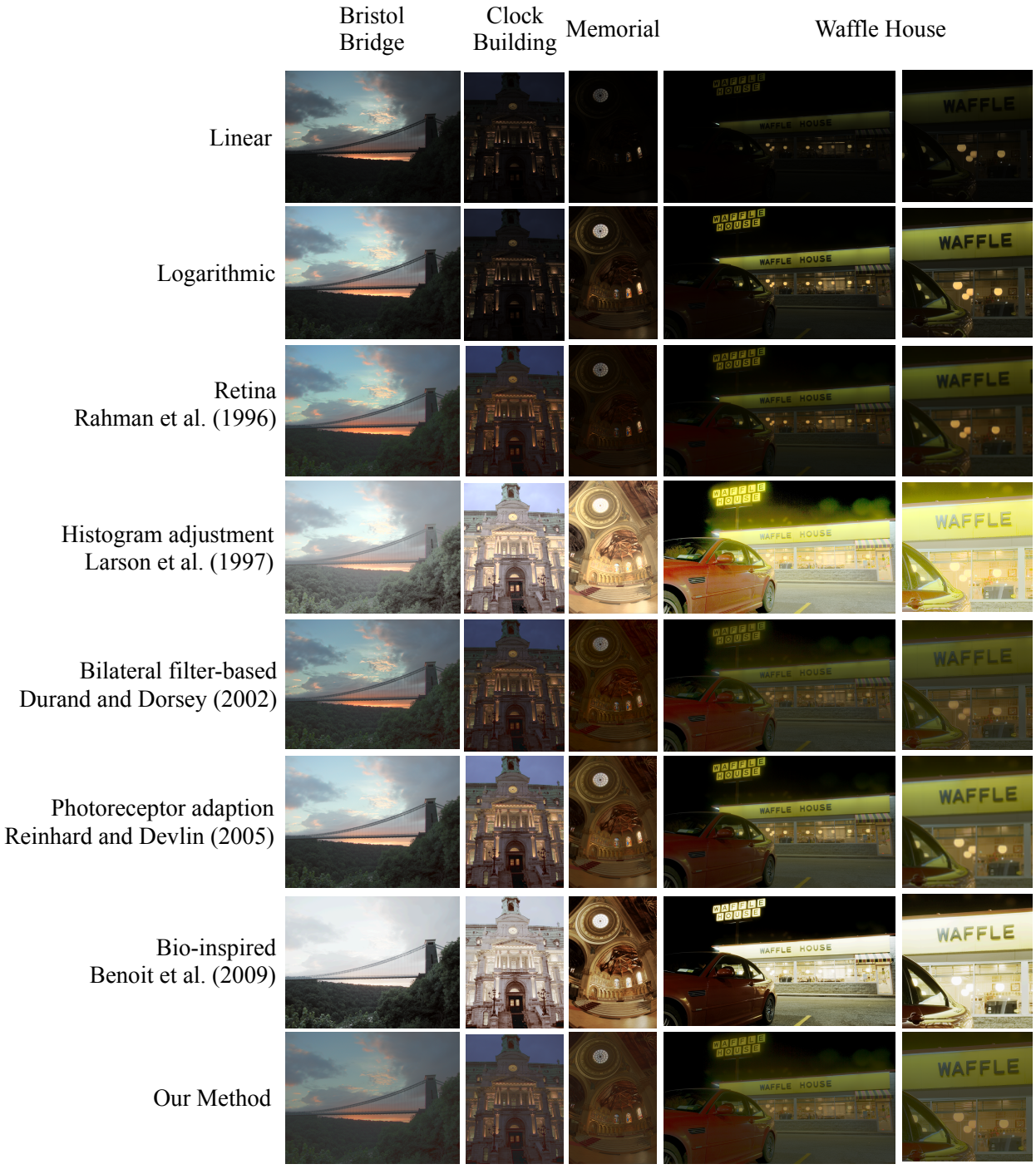


Figure 9: **Qualitative comparison between different TMOs and our method** (see text for references). Bristol Bridge and Clock Building images are copyright of Greg Ward. Memorial image is copyright of Paul Debevec. Waffle House is copyright of Mark D. Fairchild.

necessary since all the images are already calibrated. For these images the alpha factor (i.e., the key of the scene) is close to a logarithmic average of the input luminosity. As a consequence, it is not important to know if other approaches include or not this calibration step. Regarding gamma correction, for all the TMOs evaluated we checked that the output image was passed through a gamma correction layer ($\gamma = 2.2$) before being evaluated. If this was not in the original code, we added it.

In order to obtain quantitative comparisons with other TMOs, we computed the TMQI [104] for 29 different TMOs over a big subset of the HDR Photographic Survey[7], by Mark Fairchild (a total of 101 images). For each TMO we represent the distribution of the TMQI obtained for all the images in the dataset using violin plots (see Fig. 10**(a)**). Note that violin plots were chosen because they bring more information than standard box-plots. For example, they can detect bimodal behaviors while the classical box-plots miss this information. We can observe the variety of performances obtained for the different TMOs evaluated, and that the method here proposed has a comparable performance with the best methods reported in the literature. The methods evaluated, from top to bottom are: Linear, Logarithmic (see [84]), [69, 41, 65, 18, 94, 31, 98, 88, 32, 78, 52, 73, 74, 1, 26, 29, 83, 19, 24, 103, 53, 81, 48, 46, 9, 70, 108], and our method. The [108] implementation was done by us, the Kim-Krautz consistent TMO was made available by [6], Li method was provided by Yuanzhen Li, the Democratic TMO implementation was done by Magnus Oskarsson, Benoit TMO code was retrieved from the OpenCV 2.0 source package and the rest of TMO implementations were provided by [84].

Additionally, in Figures 10**(b-c)** we show a more precise comparison with the two competing bio-inspired approaches reported in Fig. 10**(a)**, namely [8] and [108]. Since TMOs yield high variability in the results depending on the input, it is of interest to evaluate how our method performs against them on each sample input individually. For both approaches we compare the TMQI value for each image of the Fairchild dataset. In Fig. 10**(b)** we compare our method with [8]. Methods seem to be complementary: there are images where our method obtained a TMQI value between 0.7 and 0.8, while [8] obtained a score over 0.8. By the contrary, images where the TMQI value of [8] did not passed over 0.8, our method generated a TMQI value in the range 0.8–1.0. Overall, our method gives a better score in 59 out of 101 images. The regression applied to this data ($r = -0.15, p = 0.13$) confirm the complementary performance showing no correlation between them. Similarly, in Fig. 10**(b)** we compare our method with [108]. Overall, our method gives a better score in 56 out of 101 images. In this case, the regression applied to data suggests a slight correlation between both TMQI values ($r = 0.33, p < 0.05$).

3.6. Result videos

Common artifacts in video TMO are brightness flickering and ghosting produced by temporal filtering. We tested our method on the two HDR videos from the Grzegorz Krawczyk dataset.[8] Results are shown in Fig. 11. Videos with the resulting radiance maps can be found in the paper website.

Our model generates satisfying results regarding radiance maps, particularly when the car driver faces direct sunlight generating overexposure in the camera. Brightness flickering is reduced as shown by the evolution of the temporal average (compare temporal variations in I_{OPL} and V_{Bip}. However, we found that this time coherence is lost if the video frames are then colorized and gamma corrected, bringing back some chromatic flickering. This is a limitation our our approach coming from the colorization method we chose [57], since flickers in the input are transmitted to the output.

Another interesting observation from Fig. 11 is about the temporal effect of CGC layer. We show the mean log luminance of each frame at different levels of our architecture: the input luminance map (L), the output of the OPL ($I_{\mathrm{OPL}t}$) and the output CGC stage (V_{Bip}). Input sequence has fast changes in the overall luminance (flickering) which are still present at the OPL stage. These temporal variations are smoothed in the next stage thanks to CGC. This effect is best seen on the videos shown on the paper website.

We found in our model that a critical parameter in terms of temporal coherence is τ_A, as seen in the $g_A(x, y, t)$ component of Eq. (9). As explained in [101], this constant relates to the time scale of the

[7] Mark Fairchild dataset website: `http://rit-mcsl.org/fairchild/HDR.html`

[8] Grzegorz Krawczyk dataset website: `http://resources.mpi-inf.mpg.de/hdr/video/`

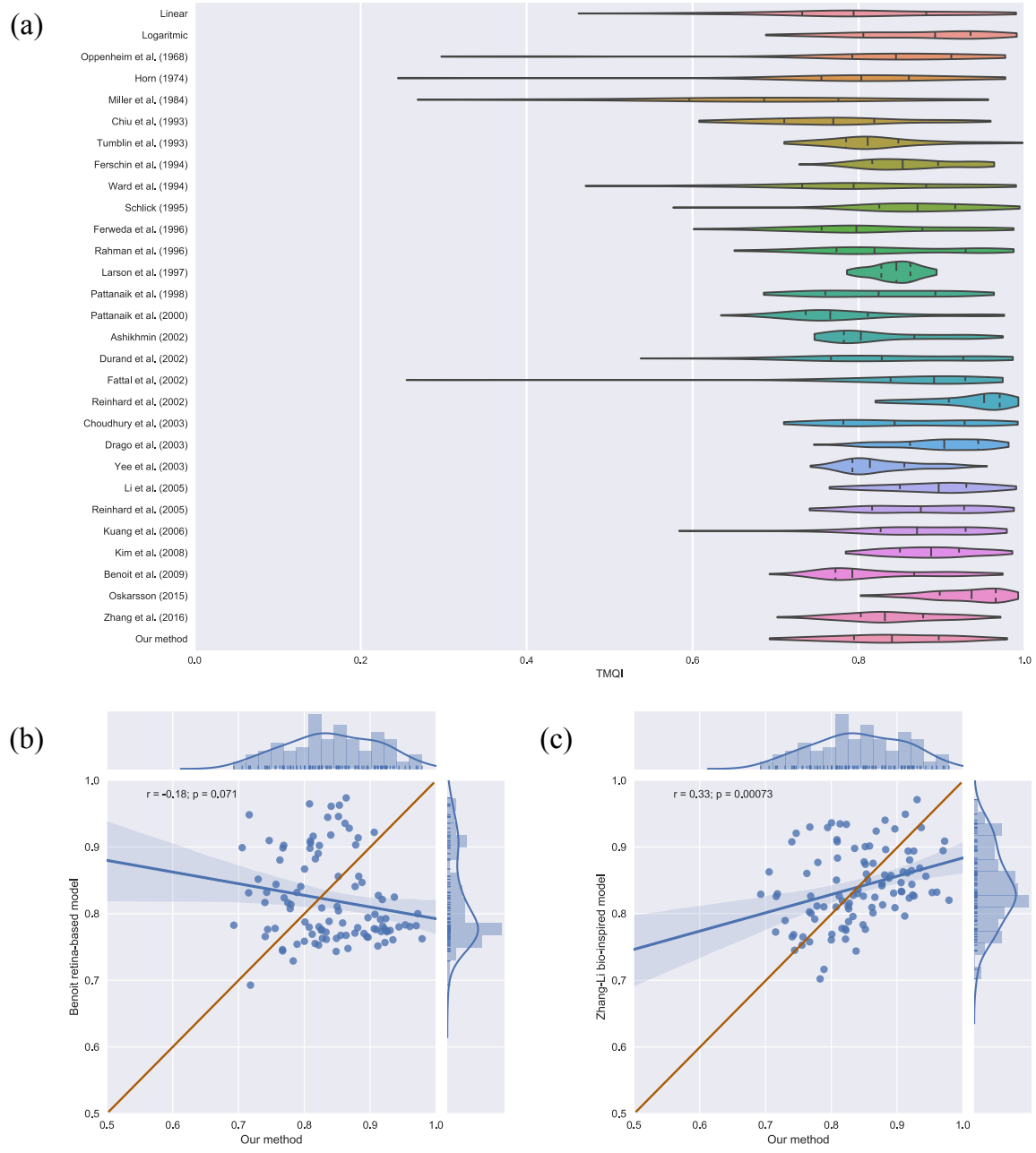


Figure 10: **Quantitative comparison between different TMOs and our method** **(a)** Comparison of the TMQI metric obtained by our method against other 29 TMOs (see text for the exact references). **(b)** Detailed comparison between TMQI values obtained by [8] and our method. We observe no correlation ($r = -0.18, p = 0.071$). **(c)** Detailed comparison between TMQI values obtained by [108] and our method. We observe a correlated response ($r = 0.33, p < 0.05$). **(b)-(c)** Orange diagonal line shows the region of equal performance. Points below this line correspond to images where our method obtained a better TMQI value.

adaptation and cannot be fixed from experimental data, because of the hypothetical nature of the stage. Since this time constant is related to a exponential kernel, and thus a low-pass time filter, it is inversely proportional to the cut off frequency of such filter. We found through tests that the best value for this is $\tau_A = 0.0005$ [s], for which temporal coherence is improved compared with the case without temporal convolutions. On the other hand, a bigger selection of τ_A leads to an strong ghosting effect, because the tight bandwidth in the filter prevents the natural changes in luminance across time per pixel, in contrast to the luminance flickering we want to avoid.

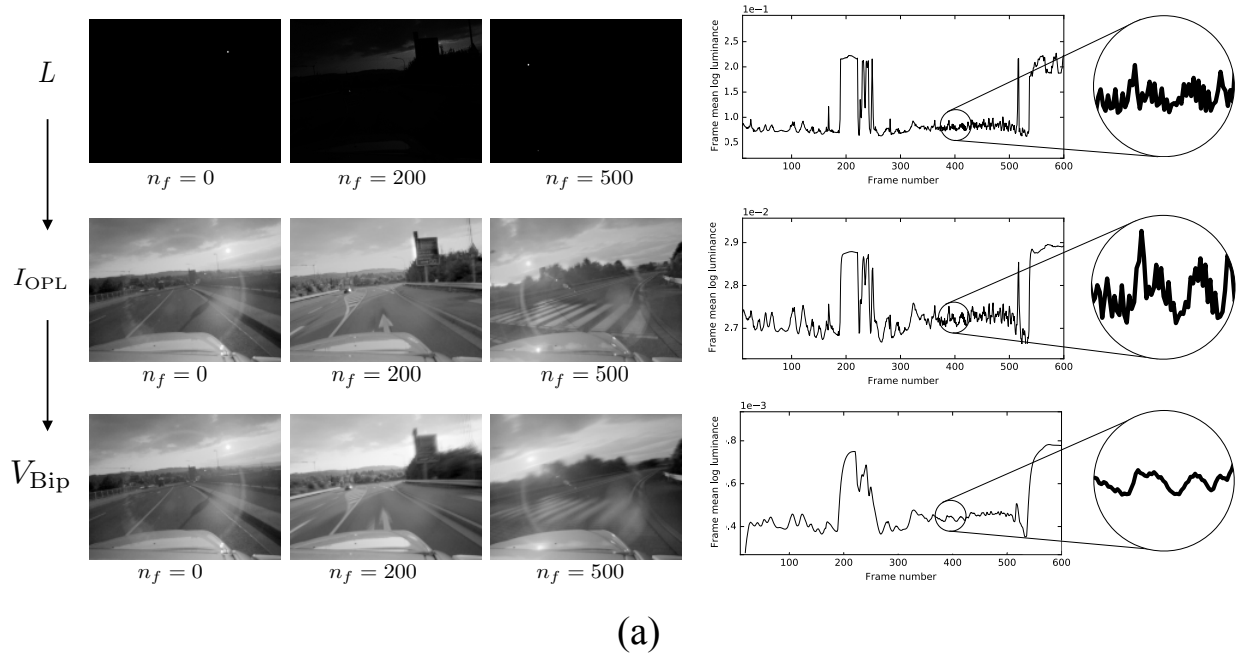
4. Conclusion

In this paper we show a promising way to address computational photography challenges by exploiting the current research in neuroscience about retina processing. Models based properties of the visual system, and more specifically on detailed retinal circuitry, will certainly benefit from the research in neuroscience. As such, choosing Virtual Retina appears to be a promising idea since our goal is to pursue its development towards a better understanding of retinal processing [14]. Indeed, one of our goal is to systematically investigate how new mechanisms found in the retina could be interpreted and be useful in the context of artificial vision problems. This conveys to this work a high potential for further improvements.

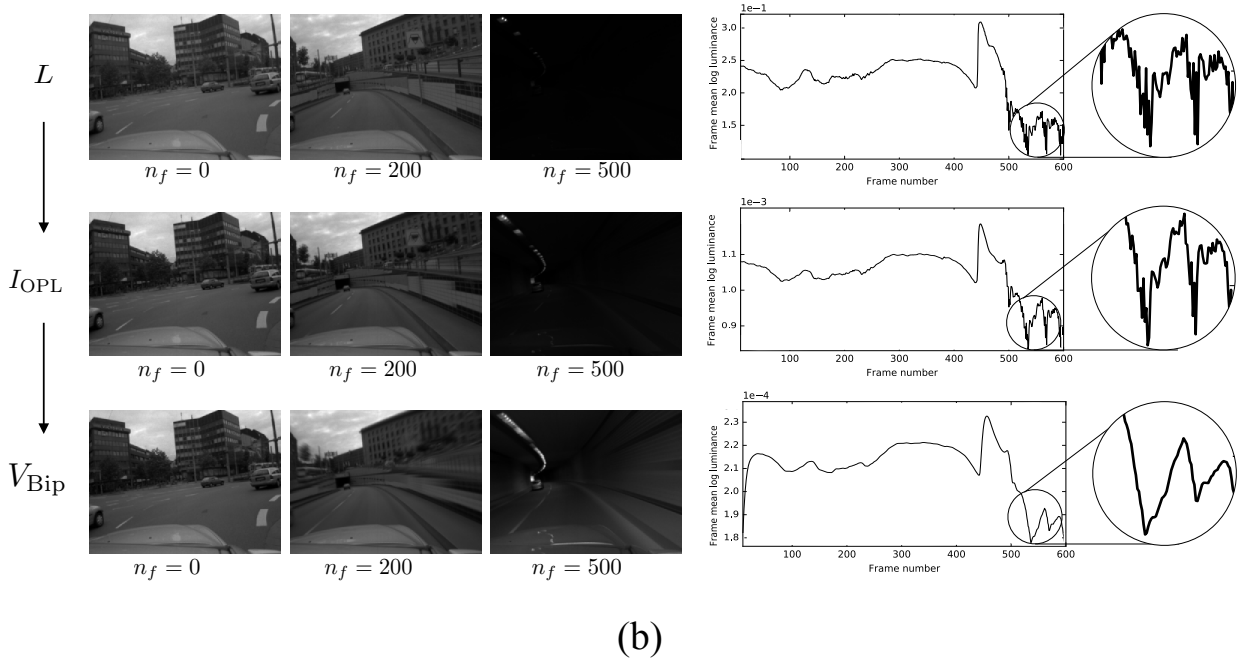
For static images our TMO works as expected, effectively compressing the dynamic range to fit a normal 24-bit LDR image generating an equalized image. In the case of videos, the TMO here proposed also work properly in the luminance information, being the CGC layer an important step to guarantee the temporal coherence to the video. In videos, one of the limitations of the TMO here presented is that the current colorization method generates color flickering on the output. This limitation arises because the original Virtual Retina model is based on luminance maps only, i.e., monochrome. A workaround of this problem could be either to include color processing in retina model or maybe improve colorization method in post-processing stage.

The model here presented contains a series of parameters related to the different stages of global and local adaptation of the input luminance image. We have observed in Section 3 the effect of different parameters in the global and local contrast adaptation. Interestingly, the parameters related to the $I_{\mathrm{OPL}}(x, y, t)$ computation relating the balance between center and surround, define the extent of local adaptation. Areas with low contrast looks brighter is the size of the surround neighborhood increases. This type of computation shows similarities with observations in neurophysiology data of sensory systems, where the system integrates information inside a wider neighborhood if the signal to noise ratio is small, or more drastically, the shape of the surround varies according to the input information [71, 33]. This behavior suggests that an adaptive selection of the surround size and shape, both at the OPL and CGC could highly benefit the quality of the TMO.

We see three challenging perspectives for this work. First perspective is to include more cells types to have a richer representation of the input and know how to exploit them. It has been shown that the retina contains a wide variety of ganglion cell types simultaneously encoding features of the visual world [35, 87, 59, 54]. A natural extension of the model here proposed could be to increase the types of retinal ganglion cells in order to expand its computational capabilities. This extension should be also joined with the proposition of efficient readouts to obtain a single retina output traduced on an output image. The readout of neural activity is related to a decoding mechanism, which is an open research topic in the neuroscience community [77, 99, 75, 36]. Second perspective is to further investigate other nonlinear mechanisms in the retina such as [10, 15, 91]. These mechanisms have been shown to account for temporal processing such as anticipation and we plan to study how they would impact the quality of tone mapped videos. Third perspective is to include fixational eye movements for the static image scenario [58]. Indeed, our eyes are never at rest. They are continuously moving even in a fixation tasks. These eye movements, and in particular the microsaccades part, are known to enhance visual perception. This idea has been explored in the context of edge detection [40], superresolution [105] and coding [62]. We think it is interesting to investigate their effect also for tone mapping. Finally, the quantitative comparison through TMQI values with other bio-inspired methods suggested complementary behaviors as we observed in Fig. 10(b)-(c). To analyze in detail the complementary contributions of each computation stage, projecting to merge the better



(a)



(b)

Figure 11: **Results on the two videos from Grzegorz Krawczyk's dataset**: (a) Highway video and (b) Tunnel video. For each case we show sample frames of the video at different levels of our architecture: the input luminance map (L), the output of the OPL (I_{OPL}) and the output CGC stage (V_{Bip}). On the right hand side, we show the time-evolution of the corresponding log spatial average of each frame. Results show how temporal coherence is improved at CGC level.

aspects of different TMOs in a single one, could be an interesting track to explore. In particular, the color de/multiplexing scheme proposed by [9] is of special interest for us to deal with videos and might reduce the flickering observed by our current colorization approach.

Appendix A. Photoreceptor adaptation and pupil model

Photoreceptors adjust their sensitivity over time and space, providing a first adaptation step for the incoming light. The VIRTUAL RETINA model proposed by [101], originally designed to process LDR images, does not have this adaptation in consideration at the input stage so there is a need to provide it without interfering the mechanisms of the rest of the model.

In the other hand, [25] showed that the photoreceptor adaptation is mutually exclusive from the adaptation occurring further into the retina, since the latter happens as the signal go from the cone bipolar cells to the ganglionar layer. Also, the authors provided a response model for cones, which in light of the previous statement, would fit well with our model.

Using electrophysiological data recorded from L type cone response, [25] constructed an empirical model which relates the normalized response of a cone $h(\cdot)$ to the background intensity i_{B} cast over it, presented in the following relation:

$$h(i_{\mathrm{B}}) = \frac{1}{1 + \left(\frac{i_{1/2}}{i_{\mathrm{B}}}\right)^{n}}, \tag{A.1}$$

where $i_{1/2}$ means the *half maximal background* intensity, estimated at $45 \cdot 10^3 \pm 7 \cdot 10^3 \left[\mathrm{P\ s^{-1}}\right]$ on their data. On the other hand, n is the Hill exponent describing the relationship between response amplitude and background intensity i_{B}, value estimated at $n = 0.7 \pm 0.05$.

The model proposed in (A.1) is expressed in *Absorbed photons per cone per second* - $\left[\mathrm{P\ s^{-1}}\right]$, and to easily use it in our model we need to convert it into $\left[\mathrm{cd\ m^{-2}}\right]$. The following formula is used to convert between these units[9]:

$$i = f(l) = 10\pi \cdot l \cdot \rho_{\mathrm{pupil}}^2, \tag{A.2}$$

where ρ_{pupil} is the radius of the pupil in millimeters and l is the luminance to convert.

There are several models to relate the radius of the pupil ρ_{pupil} with luminance covering different complexities [100]. For the sake of simplicity, we chose the [20] model defined as:

$$\rho_{\mathrm{pupil}}(\bar{L}) = 3.5875 \exp\left(-0.00092\left(7.597 + \log \bar{L}\right)^3\right), \tag{A.3}$$

which relates the average incoming spatial luminance $\bar{L}(t) = \int_{x,y} L(x,y,t)dxdy$, with the radius of the pupil ρ_{pupil} in millimeters. This model comes from an experimental fit which does not consider additional factors such as age, contraction / dilation or adapting field size.

Thus finally, the response of the photoreceptos $h(\cdot)$ can be now expressed in terms of the luminance background l_B and the background luminance $\bar{L}$ as follows:

$$h(l_{\mathrm{B}}, \bar{L}) = \frac{1}{1 + \left(\frac{i_{1/2}}{10\pi \cdot l_{\mathrm{B}} \cdot \rho_{\mathrm{pupil}}(\bar{L})^2}\right)^{n}}. \tag{A.4}$$

Furthermore, since $i_{1/2}$ is constant, we define a new expression $l_{1/2}(\bar{L})$:

$$h(l_{\mathrm{B}}, \bar{L}) = \frac{1}{1 + \left(\frac{l_{1/2}(\bar{L})}{l_{\mathrm{B}}}\right)^{n}} \tag{A.5}$$

$$l_{1/2}(\bar{L}) = \frac{i_{1/2}}{10\pi \cdot \rho_{\mathrm{pupil}}(\bar{L})^2} \tag{A.6}$$

[9] `http://retina.anatomy.upenn.edu/~rob/lance/units_photometric.html`

where $l_{1/2}$ represents the *half maximal background* luminosity for the fitted data - now in $\left[\text{cd m}^{-2}\right]$ - and l_{B}. Figure A.12 shows how $l_{1/2}$ depends on the average luminance $\bar{L}$.

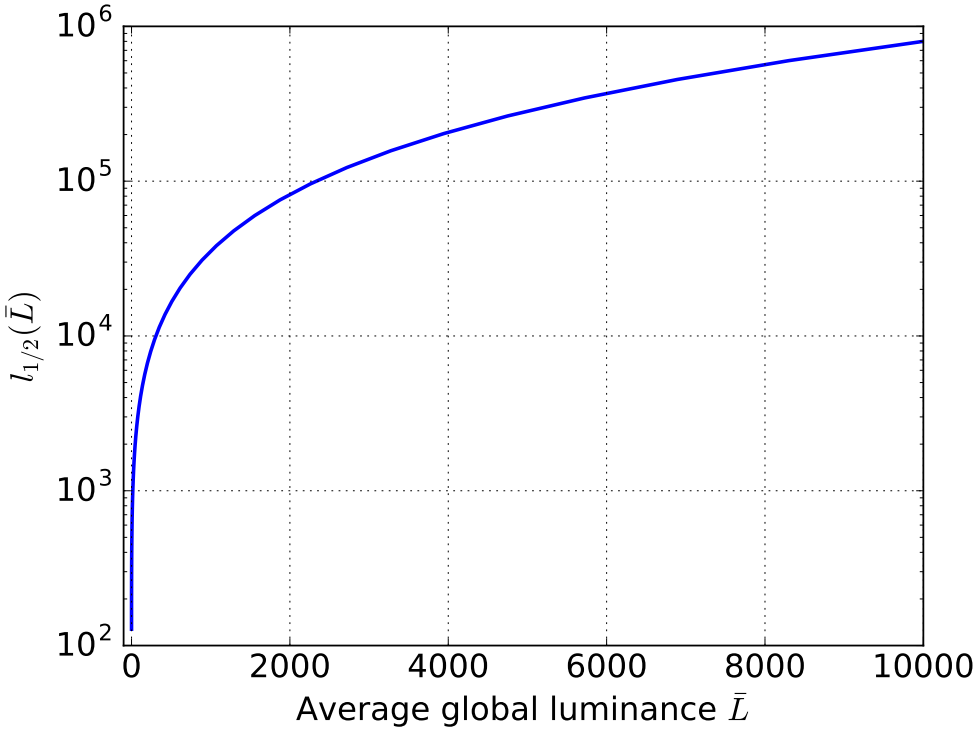


Figure A.12: **Characteristic curve of** $i_{1/2}(\bar{L})$. Plot of Equation (A.6) in function of the incoming average luminance $\bar{L}$. For each image $i_{1/2}(\bar{L})$ behaves as a constant on Eq. (A.5), yet over a video sequence the non-linear property of the curve helps to smooth luminance transients between frames.

Acknowledgements

Financial support: CONICYT-FONDECYT 1140403, CONICYT-Basal Project FB0008, Inria. The authors are very grateful to Adrien Bousseau for the fruitful discussions and his valuable feedback on several preliminary versions of this paper.

[1] Ashikhmin, M., 2002. A tone mapping algorithm for high contrast images. In: Proceedings of the 13th Eurographics Workshop on Rendering. EGRW '02. Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, pp. 145–156.

[2] Aydin, T. O., Mantiuk, R., Myszkowski, K., Seidel, H.-P., Aug. 2008. Dynamic range independent image quality assessment. ACM Transactions on Graphics (TOG) 27 (3).

[3] Aydin, T. O., Stefanoski, N., Croci, S., Gross, M., Smolic, A., 2014. Temporally coherent local tone mapping of hdr video. ACM Transactions on Graphics (TOG) 33 (6), 196.

[4] Baccus, S., Meister, M., 2002. Fast and slow contrast adaptation in retinal circuitry. Neuron 36 (5), 909–919.

[5] Balasubramanian, V., Sterling, P., 2009. Receptive fields and functional architecture in the retina. The Journal of Physiology 587 (12), 2753–2767.

[6] Banterle, F., Artusi, A., Debattista, K., Chalmers, A., 2011. Advanced High Dynamic Range Imaging: Theory and Practice. AK Peters (CRC Press), Natick, MA, USA.

[7] Basalyga, G., Montemurro, M. A., Wennekers, T., 2013. Information coding in a laminar computational model of cat primary visual cortex. J. Comput. Neurosci. 34, 273–83.

[8] Benoit, A., Alleysson, D., Hérault, J., Le Callet, P., 2009. Spatio-temporal tone mapping operator based on a retina model. In: Computational Color Imaging Workshop.

[9] Benoit, A., Caplier, A., Durette, B., Herault, J., 2010. Using human visual system modeling for bio-inspired low level image processing. Computer Vision and Image Understanding 114 (7), 758 – 773.

[10] Berry, M., Brivanlou, I., Jordan, T., Meister, M., 1999. Anticipation of moving stimuli by the retina. Nature 398 (6725), 334–338.

[11] Bertalmío, M., 2014. Image Processing for Cinema. CRC Press.
[12] Carandini, M., Demb, J. B., Mante, V., Tollhurst, D. J., Dan, Y., Olshausen, B. A., Gallant, J. L., Rust, N. C., Nov. 2005. Do we know what the early visual system does? Journal of Neuroscience 25 (46), 10577–10597.
[13] Cessac, B., Kornprobst, P., Kraria, S., Nasser, H., Pamplona, D., Portelli, G., Viéville, T., 2017. Pranas: a new platform for retinal analysis and simulation. Frontiers in NeuroinformaticsTo appear.
[14] Cessac, B., Kornprobst, P., Kraria, S., Nasser, H., Pamplona, D., Portelli, G., Viville, T., 2017. PRANAS: a new platform for retinal analysis and simulation. Frontiers in NeuroinformaticsTo appear.
[15] Chen, E. Y., Marre, O., Fisher, C., Schwartz, G., Levy, J., da Silviera, R. A., Berry, M. J. I., 2013. Alert response to motion onset in the retina. Journal of Neuroscience 33 (1), 120132.
[16] Chichilnisky, E. J., 2001. A simple white noise analysis of neuronal light responses. Network: Comput. Neural Syst. 12, 199–213.
[17] Chichilnisky, E. J., Kalmar, R. S., Apr. 2002. Functional asymmetries in ON and OFF ganglion cells of primate retina. The Journal of Neuroscience 22 (7), 27372747.
[18] Chiu, K., Herf, M., Shirley, P., Swamy, S., Wang, C., Zimmerman, K., 1993. Spatially nonuniform scaling functions for high contrast images. In: In Proceedings of Graphics Interface 93. pp. 245–253.
[19] Choudhury, P., Tumblin, J., 2003. The trilateral filter for high contrast images and meshes. In: Proceedings of the 14th Eurographics Workshop on Rendering. EGRW '03. Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, pp. 186–196.
[20] de Groot, S. G., Gebhard, J. W., 1952. Pupil size as determined by adapting luminance. J. Opt. Soc. Am. 42 (7), 492–495.
[21] Debevec, P. E., Malik, J., 1997. Recovering high dynamic range radiance maps from photographs. In: Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques. SIGGRAPH '97. ACM Press/Addison-Wesley Publishing Co., New York, NY, USA, pp. 369–378.
[22] Doutsi, E., Fillatre, L., Antonini, M., Gaulmin, J., 2015. Retinal-inspired filtering for dynamic image coding. In: Image Processing (ICIP), 2015 IEEE International Conference on. IEEE, p. 35053509.
[23] Doutsi, E., Fillatre, L., Antonini, M., Gaulmin, J., 2015. Video analysis and synthesis based on a retinal-inspired frame. In: Signal Processing Conference (EUSIPCO), 2015 23rd European. IEEE, p. 22262230.
[24] Drago, F., Myszkowski, K., Annen, T., Chiba, N., 2003. Adaptive logarithmic mapping for displaying high contrast scenes. Computer Graphics Forum 22 (3), 419–426.
[25] Dunn, F. A., Lankheet, M. J., Rieke, F., 2007. Light adaptation in cone vision involves switching between receptor and post-receptor sites. Nature 449 (7162), 603–606.
[26] Durand, F., Dorsey, J., 2002. Fast bilateral filtering for the display of high-dynamic-range images. In: Akeley, K. (Ed.), Proceedings of the SIGGRAPH. ACM Press, ACM SIGGRAPH, Addison Wesley Longman.
[27] Eilertsen, G., Mantiuk, R. K., Unger, J., 2017. A Comparative Review of Tone-mapping Algorithms for High Dynamic Range Video. Computer Graphics Forum 36 (2), 565592.
[28] Eilertsen, G., Wanat, R., Mantiuk, R., Unger, J., Oct. 2013. Evaluation of tone mapping operators for hdr-video. Computer Graphics Forum 32 (7), 275–284.
[29] Fattal, R., Lischinski, D., Werman, M., Jul. 2002. Gradient domain high dynamic range compression. ACM Trans. Graph. 21 (3), 249–256.
[30] Ferradans, S., Bertalmio, M., Provenzi, E., Caselles, V., Oct. 2011. An Analysis of Visual Adaptation and Contrast Perception for Tone Mapping. IEEE Transactions on Pattern Analysis and Machine Intelligence 33 (10), 2002–2012.
[31] Ferschin, P., Tastl, I., Purgathofer, W., Nov 1994. A comparison of techniques for the transformation of radiosity values to monitor colors. In: Proceedings of 1st International Conference on Image Processing. Vol. 3. pp. 992–996 vol.3.
[32] Ferwerda, J. A., Pattanaik, S. N., Shirley, P., Greenberg, D. P., 1996. A model of visual adaptation for realistic image synthesis. In: Proceedings of the 23rd annual conference on Computer graphics and interactive techniques. ACM, pp. 249–258.
[33] Fournier, J., Monier, C., Pananceau, M., Frgnac, Y., 2011. Adaptation of the simple or complex nature of v1 receptive fields to visual statistics. Nature neuroscience 14 (8), 10531060.
[34] Garvert, M. M., Gollisch, T., 2013. Local and global contrast adaptation in retinal ganglion cells. Neuron 77 (5), 915928.
[35] Gollisch, T., Meister, M., Jan. 2010. Eye smarter than scientists believed: neural computations in circuits of the retina. Neuron 65 (2), 150–164.
[36] Graf, A. B., Kohn, A., Jazayeri, M., Movshon, J. A., 2011. Decoding the activity of neuronal populations in macaque primary visual cortex. Nature neuroscience 14 (2), 239245.
[37] Gryaditskya, Y., Pouli, T., Reinhard, E., Seidel, H.-P., Oct. 2014. Sky based light metering for high dynamic range images. Comput. Graph. Forum 33 (7), 61–69.
[38] Hafner, D., Demetz, O., Weickert, J., 2014. Simultaneous hdr and optic flow computation. In: Pattern Recognition (ICPR), 2014 22nd International Conference on. IEEE, pp. 2065–2070.
[39] Hérault, J., 2010. Vision: Images, Signals and Neural Networks: Models of Neural Processing in Visual Perception. World Scientific.
[40] Hongler, M., de Meneses, Y. L., Beyeler, A., Jacot, J., Sep. 2003. The resonant retina: Exploiting vibration noise to optimally detect edges in an image. IEEE Transactions on Pattern Analysis and Machine Intelligence 25 (9), 1051–1062.
[41] Horn, B. K., 1974. Determining lightness from an image. Computer Graphics and Image Processing 3 (4), 277 – 299.
[42] Irawan, P., Ferwerda, J. A., Marschner, S. R., 2005. Perceptually based tone mapping of high dynamic range image streams. In: Rendering Techniques. pp. 231–242.
[43] Kalantari, N. K., Shechtman, E., Barnes, C., Darabi, S., Goldman, D. B., Sen, P., Nov. 2013. Patch-based high dynamic range video. ACM Trans. Graph. 32 (6), 202:1–202:8.

[44] Kastner, D. B., Baccus, S. A., Apr. 2014. Insights from the retina into the diverse and general computations of adaptation, detection, and prediction. Current Opinion in Neurobiology 25, 63–69.
[45] Kim, K. J., Rieke, F., Jan. 2001. Temporal contrast adaptation in the input and output signals of salamander retinal ganglion cells. Journal of Neuroscience 21 (1), 287–299.
[46] Kim, M. H., Kautz, J., 2008. Consistent tone reproduction. In: Proceedings of the Tenth IASTED International Conference on Computer Graphics and Imaging. CGIM '08. ACTA Press, Anaheim, CA, USA, pp. 152–159.
[47] Kremkow, J., Jin, J., Komban, S. J., Wang, Y., Lashgari, R., Li, X., Jansen, M., Zaidi, Q., , Alonso, J.-M., 2014. Neuronal nonlinearity explains greater visual spatial resolution for darks than lights. PNAS 111 (8), 31703175.
[48] Kuang, J., Johnson, G. M., Fairchild, M. D., 2007. icam06: A refined image appearance model for hdr image rendering. Journal of Visual Communication and Image Representation 18 (5), 406 – 414, special issue on High Dynamic Range Imaging.
[49] Kung, J., Yamaguchi, H., Liu, C., Johnson, G., Fairchild, M., Jul. 2007. Evaluating hdr rendering algorithms. ACM Transactions on Applied Perception 4 (2).
[50] Land, E. H., McCann, J. J., 1971. Lightness and retinex theory. Josa 61 (1), 1–11.
[51] Larson, G. W., Rushmeier, H., Piatko, C., 1997. A visibility matching tone reproduction operator for high dynamic range scenes. IEEE Transactions on Visualization and Computer Graphics 3 (4), 291–306.
[52] Larson, G. W., Rushmeier, H., Piatko, C., Oct 1997. A visibility matching tone reproduction operator for high dynamic range scenes. IEEE Transactions on Visualization and Computer Graphics 3 (4), 291–306.
[53] Li, Y., Sharan, L., Adelson, E. H., Jul. 2005. Compressing and companding high dynamic range images with subband architectures. ACM Trans. Graph. 24 (3), 836–844.
[54] Lorach, H., Benosman, R., Marre, O., Ieng, S.-H., Sahel, J. A., Picaud, S., Oct. 2012. Artificial retina: the multichannel processing of the mammalian retina achieved with a neuromorphic asynchronous light acquisition device. Journal of Neural Engineering 9 (6), 066004.
[55] Mangiat, S., Gibson, J., 2010. High dynamic range video with ghost removal. In: SPIE Optical Engineering+ Applications. International Society for Optics and Photonics, pp. 779812–779812.
[56] Mantiuk, R., Daly, S., Kerofsky, L., 2008. Display adaptive tone mapping. In: ACM Transactions on Graphics (TOG). Vol. 27. ACM, p. 68.
[57] Mantiuk, R., Tomaszewska, A., Heidrich, W., 2009. Color correction for tone mapping. In: Computer Graphics Forum. Vol. 28. Wiley Online Library, pp. 193–202.
[58] Martinez-Conde, S., Otero-Millan, J., Macknik, S. L., Feb. 2013. The impact of microsaccades on vision: towards a unified theory of saccadic function. Nature Reviews Neuroscience 14 (2), 83–96.
[59] Masland, R. H., Jun. 2011. Cell populations of the retina: The proctor lecture. Investigative Ophthalmology and Visual Science 52 (7), 4581–4591.
[60] Masmoudi, K., Antonini, M., Kornprobst, P., 2010. Another look at the retina as an image scalar quantizer. In: Proceedings of the International Symposium on Circuits and Systems (ISCAS).
[61] Masquelier, T., 2012. Relative spike time coding and stdp-based orientation selectivity in the early visual system in natural continuous and saccadic vision: a computational model. Journal of Computational Neuroscience 32 (3), 425–441.
[62] Masquelier, T., Portelli, G., Kornprobst, P., Apr. 2016. Microsaccades enable efficient synchrony-based coding in the retina: a simulation study. Scientific Reports 6, 24086.
[63] Medathati, N. V. K., Neumann, H., Masson, G. S., Kornprobst, P., 2016. Bio-inspired computer vision: Towards a synergistic approach of artificial and biological vision. Computer Vision and Image Understanding, –.
[64] Meylan, L., Alleysson, D., Süsstrunk, S., 2007. Model of retinal local adaptation for the tone mapping of color filter array images. J. Opt. Soc. Am. A 24 (9), 2807–2816.
[65] Miller, N. J., Ngai, P. Y., Miller, D. D., 1984. The application of computer graphics in lighting design. Journal of the Illuminating Engineering Society 14 (1), 6–26.
[66] Mohemmed, A., Lu, G., Kasabov, N., 2012. Evaluating SPAN Incremental Learning for Handwritten Digit Recognition. In: Neural Information Processing. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 670–677.
[67] Muchungi, K., Casey, M., Sep. 2012. Simulating Light Adaptation in the Retina with Rod-Cone Coupling. In: ICANN. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 339–346.
[68] O'Carroll, D. C., Warrant, E. J., Feb. 2017. Introduction: Vision in dim light: highlights and challenges. Phil. Trans. R. Soc. B.
[69] Oppenheim, A., Schafer, R., Stockham, T., Sep 1968. Nonlinear filtering of multiplied and convolved signals. IEEE Transactions on Audio and Electroacoustics 16 (3), 437–466.
[70] Oskarsson, M., 2015. Democratic Tone Mapping Using Optimal K-means Clustering. Springer International Publishing, Cham, pp. 354–365.
[71] Pack, C., Hunter, J., Born, R., 2005. Contrast dependence of suppressive influences in cortical area MT of alert macaque. Journal of Neurophysiology 93 (3), 1809–1815.
[72] Pandarinath, C., Victor, J., Nirenberg, S., 2010. Symmetry Breakdown in the ON and OFF Pathways of the Retina at Night: Functional Implications. The Journal of neuroscience 30 (30), 1000610014.
[73] Pattanaik, S. N., Ferwerda, J. A., Fairchild, M. D., Greenberg, D. P., 1998. A multiscale model of adaptation and spatial vision for realistic image display. In: Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques. SIGGRAPH '98. ACM, New York, NY, USA, pp. 287–298.
[74] Pattanaik, S. N., Tumblin, J., Yee, H., Greenberg, D. P., 2000. Time-dependent visual adaptation for fast realistic image display. In: Proceedings of the 27th annual conference on Computer graphics and interactive techniques. ACM Press/Addison-Wesley Publishing Co., pp. 47–54.

[75] Pitkow, X., Meister, M., 2012. Decorrelation and efficient coding by retinal ganglion cells. Nature neuroscience 15 (4), 628635.
[76] Purves, D., Augustine, G. J., Fitzpatrick, D., Katz, L. C., LaMantia, A.-S., McNamara, J. O., Williams, S. M., 2001. Neuroscience, 2nd Edition. Sinauer Associates, Inc.
[77] Quiroga, R. Q., Panzeri, S., 2009. Extracting information from neuronal populations: information theory and decoding approaches. Nature Reviews Neuroscience 10 (3), 173185.
[78] Rahman, Z.-u., Jobson, D. J., Woodell, G. A., 1996. A multiscale retinex for color rendition and dynamic range compression. Tech. rep., NASA Langley.
[79] Ratliff, C. P., Borghuis, B. G., Kao, Y.-H., Sterling, P., Balasubramanian, V., Oct. 2010. Retina is structured to process an excess of darkness in natural scenes. Proc Natl Acad Sci U S A. 107 (40), 1736817373.
[80] Reinhard, E., Nov. 2002. Parameter estimation for photographic tone reproduction. J. Graph. Tools 7 (1), 45–52.
[81] Reinhard, E., Devlin, K., 2005. Dynamic range reduction inspired by photoreceptor physiology. IEEE Transactions on Visualization and Computer Graphics 11 (1), 13–24.
[82] Reinhard, E., Pouli, T., Kunkel, T., Long, B., Ballestad, A., Damberg, G., Nov. 2012. Calibrated image appearance reproduction. ACM Trans. Graph. 31 (6), 201:1–201:11.
[83] Reinhard, E., Stark, M., Shirley, P., Ferwerda, J., Jul. 2002. Photographic tone reproduction for digital images. ACM Trans. Graph. 21 (3), 267–276.
[84] Reinhard, E., Ward, G., Pattanaik, S., Debevec, P., 2005. High Dynamic Range Imaging: Acquisition, Display, and Image-Based Lighting. The Morgan Kaufmann Series in Computer Graphics. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.
[85] Rieke, F., Dec. 2001. Temporal contrast adaptation in salamander bipolar cells. Journal of Neuroscience 21 (23), 9445–9454.
[86] Rieke, F., Rudd, M. E., Dec. 2009. The challenges natural images pose for visual adaptation. Neuron 64 (5), 605–616.
[87] Sanes, J., Masland, R., 2015. The types of retinal ganglion cells: current status and implications for neuronal classification. Annu Rev Neurosci 38, 221–246.
[88] Schlick, C., 1995. Quantization Techniques for Visualization of High Dynamic Range Pictures. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 7–20.
[89] Shapley, R., Enroth-Cugell, C., 1984. Visual adaptation and retinal gain controls. Progress in retinal research 3, 263–346.
[90] Shapley, R. M., Victor, J. D., 1978. The effect of contrast on the transfer properties of cat retinal ganglion cells. The Journal of Physiology 285 (1), 275–298.
[91] Souihel, S., Cessac, B., Sep. 2017. Modifying a biologically inspired retina simulator to reconstruct realistic responses to moving stimuli. Bernstein Conference 2017, submitted.
[92] Takeshita, D., Smeds, L., Ala-Laurila, P., Feb. 2017. Processing of single-photon responses in the mammalian on and off retinal pathways at the sensitivity limit of vision. Phil. Trans. R. Soc. B.
[93] Thoreson, W., Mangel, S., 2012. Lateral interactions in the outer retina. Prog Retin Eye Res. 31 (5), 407–441.
[94] Tumblin, J., Rushmeier, H., Nov 1993. Tone reproduction for realistic images. IEEE Computer Graphics and Applications 13 (6), 42–48.
[95] Van Hateren, J., 2006. Encoding of high dynamic range video with a model of human cones. ACM Transactions on Graphics (TOG) 25 (4), 1380–1399.
[96] Vance, P., Coleman, S. A., Kerr, D., Das, G., McGinnity, T., 2015. Modelling of a retinal ganglion cell with simple spiking models. In: IEEE Int. Jt. Conf. Neural Networks. pp. 1–8.
[97] Victor, J. D., 1987. The dynamics of the cat retinal X cell centre. The Journal of Physiology 386 (1), 219–246.
[98] Ward, G., 1994. Graphics gems iv. ACM, New York, NY, USA, Ch. A Contrast-based Scalefactor for Luminance Display, pp. 415–421.
[99] Warland, D., Reinagel, P., Meister, M., 1997. Decoding visual information from a population of retinal ganglion cells. J. Neurophysiol 78, 23362350.
[100] Watson, A. B., Yellott, J. I., 2012. A unified formula for light-adapted pupil size. Journal of Vision 12 (10), 12.
[101] Wohrer, A., Kornprobst, P., 2009. Virtual Retina : A biological retina model and simulator, with contrast gain control. Journal of Computational Neuroscience 26 (2), 219, dOI 10.1007/s10827-008-0108-4.
[102] Wohrer, A., Kornprobst, P., Antonini, M., Jun. 2009. Retinal filtering and image reconstruction. Tech. Rep. 6960, INRIA.
[103] Yee, Y. H., Pattanaik, S., 2003. Segmentation and adaptive assimilation for detail-preserving display for high-dynamic range images. Visual Computer 19, 457–466.
[104] Yeganeh, H., Wang, Z., Feb 2013. Objective quality assessment of tone-mapped images. IEEE Transactions on Image Processing 22 (2), 657–667.
[105] Yi, D., Jiang, P., Mallen, E., Wang, X., Zhu, J., Aug. 2011. Enhancement of image luminance resolution by imposing random jitter. Neural Computing and Applications 20 (2), 261–272.
[106] Zaghloul, K., Boahen, K., Demb, J. B., Apr. 2003. Different circuits for on and off retinal ganglion cells cause different contrast sensitivities. The Journal of Neuroscience : The Official Journal of the Society for Neuroscience 23 (7), 264554.
[107] Zhang, X.-S., Gao, S.-B., Li, C.-Y., Li, Y.-J., 2015. A retina inspired model for enhancing visibility of hazy images. Frontiers in Computational Neuroscience 9, 151.
[108] Zhang, X.-S., Li, Y.-J., 2016. A Retina Inspired Model for High Dynamic Range Image Rendering. Springer International Publishing, Cham, pp. 68–79.