

Model of retinal local adaptation for the tone mapping of color filter array images

Laurence Meylan,^{1,*} David Alleysson,² and Sabine Süsstrunk¹

¹*School of Computer and Communication Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland*

²*Psychology and NeuroCognition Laboratory, CNRS UMR 5105, Université Pierre-Mendes France (UPMF), Grenoble, France*

*Corresponding author: laurence.meylan@gmail.com

Received January 24, 2007; revised April 20, 2007; accepted April 23, 2007;
posted April 30, 2007 (Doc. ID 79299); published August 8, 2007

We present a tone mapping algorithm that is derived from a model of retinal processing. Our approach has two major improvements over existing methods. First, tone mapping is applied directly on the mosaic image captured by the sensor, analogous to the human visual system that applies a nonlinearity to the chromatic responses captured by the cone mosaic. This reduces the number of necessary operations by a factor 3. Second, we introduce a variation of the center/surround class of local tone mapping algorithms, which are known to increase the local contrast of images but tend to create artifacts. Our method gives a good improvement in contrast while avoiding halos and maintaining good global appearance. Like traditional center/surround algorithms, our method uses a weighted average of surrounding pixel values. Instead of being used directly, the weighted average serves as a variable in the Naka–Rushton equation, which models the photoreceptors' non-linearity. Our algorithm provides pleasing results on various images with different scene content and dynamic range. © 2007 Optical Society of America

OCIS codes: 100.2000, 100.2980, 110.5200, 330.1690, 330.4060, 330.6110.

1. INTRODUCTION

Most of today's digital cameras are composed of a single sensor with a color filter array (CFA) placed in front to select the spectral band that is captured at each spatial position, called a pixel (Fig. 1, left; also see [1]). Since only one chromatic component is retained for each pixel, a color reconstruction must be performed to obtain the full resolution color image with three chromatic components per pixel.

In traditional color processing workflows [2], this color reconstruction, or demosaicing [Fig. 2(a)] usually takes place before applying any rendering operations. The mosaiced image captured by the CFA is first demosaiced to obtain an RGB image with three chromatic components per spatial location. Color rendering operations, which include white balancing, color matricing, and tone mapping, are performed later.

Instead of the workflow shown in Fig. 2(a), we propose a solution where the demosaicing is the last step of the color processing workflow. Color rendering operations are thus performed directly on the CFA image [Fig. 2(b)]. In this article, we consider only the tone mapping operation of color rendering. However, color matricing and white balancing can also be implemented before demosaicing.

Our motivations to use such a workflow is that it is more analogous to the retinal processing of the human visual system (HVS) [3–5], as discussed in Section 2. Another motivation is that applying the tone mapping directly to the CFA image requires only one third of the operations. This, in addition to the use of small filters, makes our method relatively fast compared with other ex-

isting local tone mapping algorithms. Finally, because the rendering operations are performed directly on the values captured by the sensor, there is no loss of information prior to rendering.

Our tone mapping algorithm takes inspiration from the nonlinear adaptation that occurs in the retina, which efficiently improves local contrasts while conserving good global appearance [6,7]. Figures 2(c) and 2(d) show an example of applying our method to a high-dynamic-range image (i.e., containing high contrast and important image details in dark and bright areas). The left image shows the result obtained with standard global tone mapping [8,9] (in this case a gamma operator), and the right image shows the result obtained with our algorithm. Our method successfully enhances detail visibility in the center of the image, the details are well rendered without requiring an additional sharpening operation.

We applied our algorithm to various kinds of captured scenes having different dynamic ranges and different keys. Dynamic range is defined as the luminance ratio of the brightest and darkest object in the scene. High and low key are terms used to describe images that have a higher-than-average and lower-than-average mean intensity, respectively. Unlike other methods that work well only with certain kinds of images, the results show that our tone mapping operator successfully improves image appearance in all cases without creating artifacts.

This article is structured as follows: Section 2 provides background knowledge on tone mapping and the model of retinal adaptation on which we base our method. Section 3 presents the algorithm. Section 4 shows the results ob-

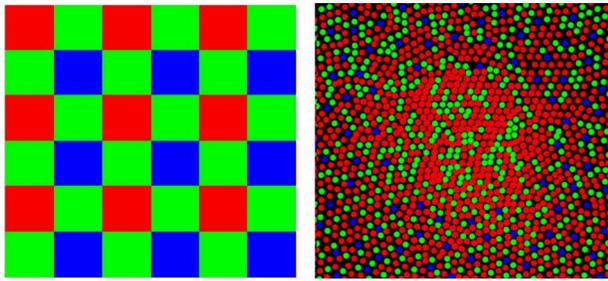


Fig. 1. (Color online) Bayer CFA (left) and the spatiochromatic sampling of the cone mosaic (right) (Inspired from Roorda *et al.* [1]).

tained by our proposed workflow, and Section 5 discusses the differences of our algorithm compared with other existing methods. Section 6 concludes the article.

2. BACKGROUND

In this section, we discuss the correspondence of our tone mapping algorithm with a simplified model of retinal processing. For this purpose, we take into consideration the sampling of chromatic information by the cone mosaic and the nonlinearity that applies to that mosaic. We concentrate on one specific nonlinear processing model proposed by Naka and Rushton [6,10] that we use in our algorithm. We discuss the properties of the CFA images on which we apply our tone mapping. Finally, tone mapping

operators in general, specifically the center/surround family of local tone mapping algorithms is also reviewed, as our method bears some similarity to the latter.

A. Model of Retinal Processing

Historically, many analogies with the HVS have been exploited to develop image and computer vision applications. For example, there is a correspondence between trichromacy (the ability of human vision to distinguish different colors given by the interaction of three kinds of photoreceptors) and the three color channels that constitute a color image [11,12]. Another equivalence exists between the spatiochromatic sampling of the cone mosaic and the sampling of color in a single-chip sensor such as given by the Bayer CFA (Fig. 1) [13,14].

Our proposed workflow [Fig. 2(b)] exploits another analogy with human vision, namely between the tone mapping operations in the image processing workflow and the nonlinear adaptation taking place in the retina. The goal here is not to precisely model the dynamics of retinal processing, such as is done, for example, by Van Hateren [15]. We aim to identify, and simplify, which type of processing acts on the retinal signal in order to develop algorithms suitable for in-camera processing. We focus on the nonlinearities applied to the mosaic of chromatic responses captured by the cones.

One role of tone mapping is to nonlinearly process the captured image to mimic the retina's nonlinear adapta-

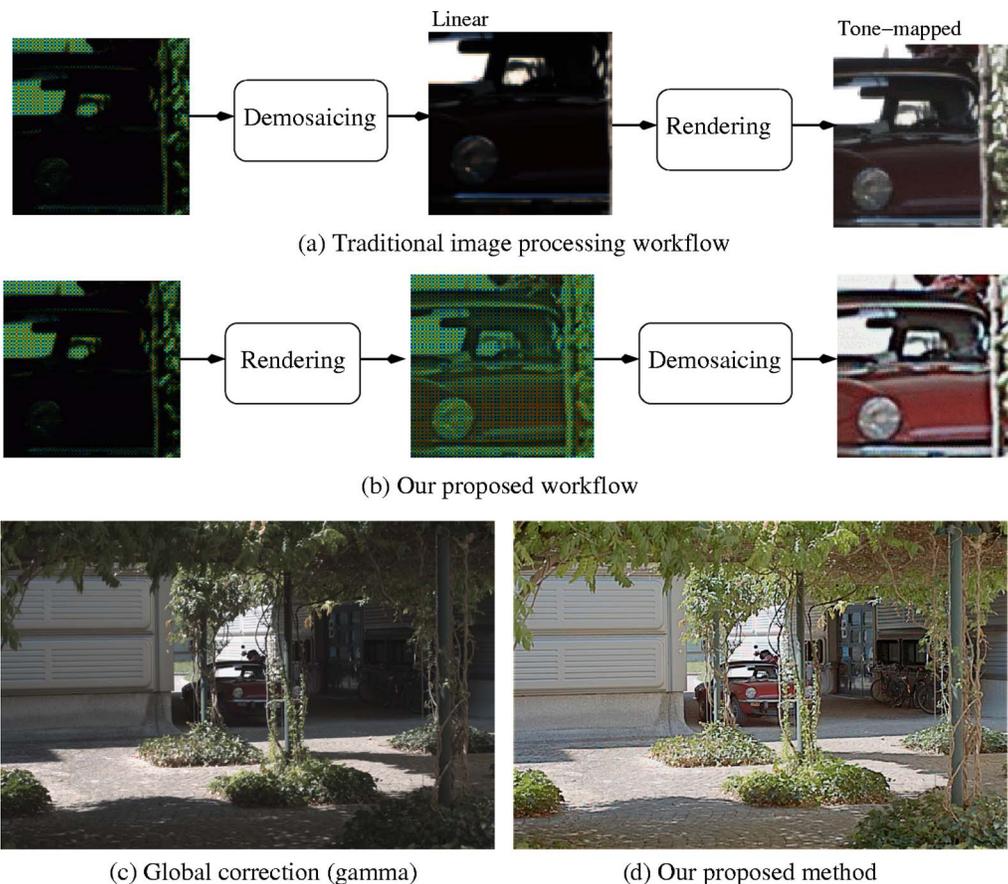


Fig. 2. (Color online) (a) Traditional image processing workflow. (b) Our proposed workflow. (c) Image rendered with a global tone mapping operator (gamma). (d) Image rendered according to our method.

tion and render the image as if the HVS had processed it. In traditional workflows, this nonlinear encoding is usually applied to the RGB color image after the color mosaic captured by the CFA sensor is demosaiced. For the HVS, the nonlinear adaptation takes place in the retina directly after light absorption by the cones. At this level, the retinal image is a spatial multiplexing of chromatic cone responses, and there is no reconstruction of full color information at each spatial position. We know that the sampled color responses are still in a mosaic representation at the output of the retina, as illustrated by the behavior of ganglion cell receptive fields [3] (see Fig. 3). We thus propose what we believe to be a new image processing workflow where the nonlinear encoding (tone mapping) is performed directly on the mosaic image provided by the Bayer CFA pattern.

Figure 3 shows the model of the retinal cell layers on which we base our algorithm (readers not familiar with the HVS can consult the web pages of Webvision [16]). We exploit the fact that the retina is composed of two functional layers, the outer plexiform layer (OPL) and the inner plexiform layer (IPL), that both apply an adaptive nonlinearity on the input signal. These two layers are composed of the cones, the horizontal and amacrine cells, which provide the horizontal connectivity, and the bipolar and ganglion cells. When the light enters the retina, it is sampled by the cones into a mosaic of chromatic components. The horizontal cells measure the spatial average of several cone responses, which determines the cones' adaptation factors through a feedback loop [17]. The color signals are then passed through the bipolar cells to the ganglion cells. We assume that the role of the bipolar cells is simply to pass the color signal from the OPL to the IPL. In the IPL, a similar nonlinear processing is applied. We assume that the amacrine cells also provide feedback to modulate the adaptive nonlinearity of the ganglion cells. This second nonlinearity has been found to provide psychophysical [18,19] and physiological [10] evidence for an adaptation mechanism to contrast rather than to intensity. Moreover, it has been suggested that this nonlinearity is postreceptoral and applies to color opponent repre-

sentation [7,19]. We assume here that it originates in the interaction among bipolar, amacrine, and ganglion cells.

Our tone mapping algorithm also applies two nonlinear processings on the CFA image in imitation of the IPL and OPL functionalities. Both non-linear operations are based on Naka and Rushton [6,10], who developed a model for the photoreceptor nonlinearities and adaptation to incoming light. Spitzer *et al.* [20] also proposed a biological model for color contrast, which used similar adaptation mechanisms. The nonlinear RGB mosaic image is then demosaiced to reconstruct the RGB tone mapped image.

B. Adaptive Nonlinearity

Our model of the OPL and IPL nonlinearities takes inspiration from the Naka–Rushton equation [6,10]

$$Y = \frac{X}{X + X_0}, \tag{1}$$

where X represents the input light intensity, X_0 is the adaptation factor, and Y is the adapted signal. In the original formulation [6], the adaptation factor (X_0) is determined by the average light reaching the entire field of view. In our method, X_0 varies for each pixel. It is a local variable given by the average light intensity in the neighborhood of one pixel. Figure 4 illustrates the Naka–Rushton function for different values of X_0 . If X_0 is small, the cell output has increased sensitivity. If X_0 is large, there is not much change in sensitivity.

In our model, the Naka–Rushton equation is used to calculate the nonlinearities of both the OPL and IPL; X_0 is given by the output of the horizontal cells or amacrine cells, respectively, and modulates the sensitivities of the cones and of the ganglion cells.

Usually, the first retinal nonlinearity is assumed to be due only to the dynamics of the photoreceptors themselves [15]. We make the hypothesis that the horizontal cell network intervenes in the light regulation of the photoreceptors. Because of its local spatial averaging characteristics, the network could allow for a more powerful regulation of the cone sensitivities. Also, horizontal cells influence the cone responses through feedback or direct

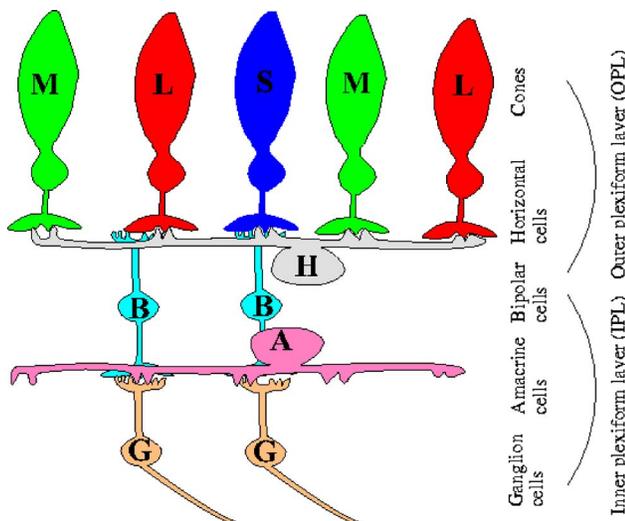


Fig. 3. (Color online) Simplified model of the retina.

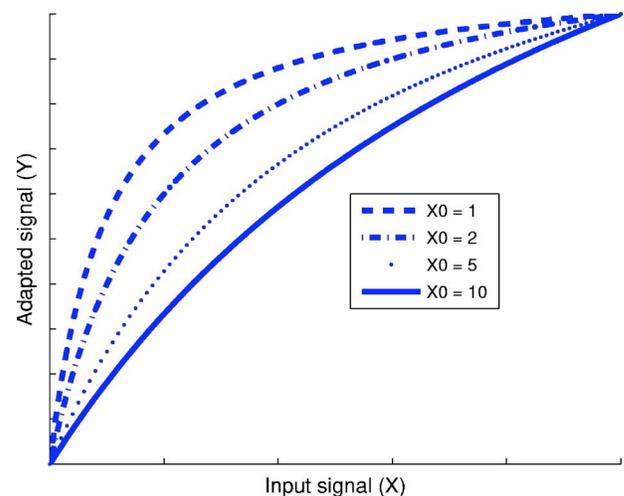


Fig. 4. Naka–Rushton function with different adaptation factors X_0 .

feedforward on bipolar cells [17]. Thus, our assumption is that the mechanism by which horizontal cells modify cone responses is due to a regulation of the cone's nonlinear adaptation factor, based on the response of the horizontal cells network at the cone location.

C. Properties of a CFA Image

The two nonlinearities described above are applied directly to the CFA image. In our implementation, the CFA image is obtained using a Bayer pattern [14] in front of the camera sensor, which results in a spatiochromatic sampling of the scene. This mosaic image has certain properties that allow the luminance and the chrominance of the image to be treated separately.

Alleysson *et al.* [21] showed that if we analyze the amplitude Fourier spectrum of a Bayer CFA image, the luminance is located in the center of the spectrum and the chrominance is located at the borders. The luminance is present at full resolution, while the chrominance is downsampled and encoded with opponent colors. It follows that a wide-band low-pass filter can be used to recover the luminance and that a high-pass or bandpass filter can recover the downsampled chrominance. Choosing the appropriate filters allows one to implement an efficient demosaicing algorithm. Their method was refined by Dubois [22] and Lian *et al.* [23], who propose a more accurate estimation of the luminance.

In Subsection 3.C, we will apply the Alleysson *et al.* method for demosaicing. In Subsections 3.A and 3.B, we use the property of localized luminance and chrominance when computing the response of the horizontal and amacrine cells as a guarantee that using a low-pass filter will indeed provide the average of the luminance in a surrounding area. In other words, we apply the nonlinearities only to the luminance signals, not to any chromatic components.

D. Tone Mapping

Tone mapping is the operation in the image processing workflow that matches scene to display luminances. The goal of tone mapping may vary, but the intent often is to reproduce visually pleasing images that correspond to the expectation of the observer.

Tone mapping algorithms can either be global (spatially invariant) or local (spatially variant). A global tone mapping is a function that maps the input pixel value to a display value, not taking into account the spatial position of the treated pixel (one input value corresponds to one and only one output value). A typical tone mapping function can be logarithmic, a power law (often referred to as a "gamma" function) or a sigmoid, also called an "s shape." More sophisticated global tone mapping methods vary the function parameters depending on global characteristics of the image [8,9,24,25]. The key of the image can be used to determine the exponent of the gamma function [24]. In Braun and Fairchild [8] and in Holm [9], an s-shaped function is defined by the image statistics, such as the mean and the variance of the intensity. In Ward *et al.* [25], the histogram distribution is used to construct an image-dependent global function.

With local tone mapping algorithms, one input pixel value can lead to different output values depending on the

pixel's surround. A local tone mapping operator is used when it is necessary to change local features in the image, such as increasing the local contrast to improve detail visibility. Many local tone mapping algorithms have been proposed, which can be grouped into different classes sharing the same common features (see Delvin [26] and Reinhard *et al.* [27] for a review). Center/surround methods take inspiration from the HVS receptive fields and lateral inhibition. They increase the local contrast by taking the difference between pixel values and an average of their surround [24,28–30]. Their common drawbacks are the creation of halos along high-contrast edges and graying out of low-contrast areas. Because center/surround methods share similarities with the proposed method, they are described in more detail in Subsection 2.E. Gradient-based methods [31] work directly on the image gradient to increase the local contrast by weighting high and low gradient values differently dependent on surrounding image data. One difficulty of this technique is integrating the gradient to recover the treated image. Frequency-based methods [32] separate the low- and high-frequency bands of the image. The low-frequency band is assumed to correspond approximatively to the illuminant and is compressed while the image details given by the high-frequency bands are kept. These techniques work well for high-dynamic-range images but are less appropriate for low-dynamic-range images.

Which tone mapping operation should be performed depends on the dynamic range of the scene. However, it also depends on the dynamic range of the display, which is given by the ratio between the brightest and the darkest display luminance (determined by the display technology and viewing conditions). In the case of a low-dynamic-range scene (e.g., a foggy scene with no high contrast), the input image's dynamic range is smaller than that of the display and thus needs to be expanded. In the opposite case of a high-dynamic-range scene (e.g., a sunset), whose dynamic range exceeds that of the display, the luminance ratio must be compressed. Since compressing high-dynamic-range images causes a loss of detail visibility over the whole tonal range, it is often necessary to apply a local tone mapping in addition to the global compression in order to increase the local contrast and keep detail visibility.

E. Center/Surround Methods

Traditional center/surround algorithms compute the treated pixel values by taking the difference in the log domain between each pixel value and a weighted average of the pixel values in its surround:

$$I'(p) = \log(I(p)) - \log(I(p) * G), \quad (2)$$

where p is a pixel in the image, I is the treated image, $*$ denotes the convolution operation, and G is a low-pass filter (often a Gaussian).

A common drawback of center/surround methods is that the increase in local contrast depends greatly on the size of the filter. When a small filter is used, halo artifacts appearing as shadows along high-contrast edges can become visible. When a large filter is used, the increase in local contrast is not sufficient to retrieve detail visibility in dark or bright areas. Another drawback of center/

surround methods is that they tend to gray out (or wash out) low-contrast areas. For example, a plain black area or a bright low-contrast zone will tend to become gray due to the local averaging.

These drawbacks have already been discussed in the literature [29,30,33], and solutions to overcome them were developed. Rahman *et al.* [30] introduced a multi-scale method where the center/surround operation is performed for three different scales so that halo artifacts and graying out are reduced. However, these artifacts are still visible when the scene contains very high contrasts. Meylan and Süsstrunk [29] introduced an adaptive filter, whose shape follows the high-contrast edges in the image and thus prevents halo artifacts. The graying out is avoided by using a sigmoid weighting function to conserve black and white low-contrast areas. Their method well retrieves details in dark areas but tends to compress highlights too much. It is also computationally very expensive, as the filter has to be recomputed for every pixel. We will compare our algorithm with these two methods in Section 4.

In general, existing center/surround tone mapping operators work well only for a limited set of images. The advantage of the algorithm presented here is that it provides a pleasing, artifact-free reproduction for all kinds of scenes (see Section 4). It can be considered to belong to the center/surround family of local tone mapping operators where the surround is used to modulate an adaptive nonlinear function rather than as a fixed factor subtracted from the input pixel.

3. LOCAL TONE MAPPING ALGORITHM FOR CFA IMAGES

Our local tone mapping method processes the images according to the retinal model that was described in Subsection 2.A. The input mosaic image (or CFA image), which has one chromatic component per spatial location, is treated by two consecutive nonlinear operations. Demosaicing is applied last in order to obtain a color image with three color components per pixel. Each of these steps is described in the following sections.

A. First Nonlinearity

The first nonlinear operation simulates the adaptive nonlinearity of the OPL. The adaptation factors, which correspond to the horizontal cell responses, are computed for each pixel by performing a low-pass filter on the input CFA image:

$$H(p) = I_{CFA}(p) * G_H + \frac{\overline{I_{CFA}}}{2}, \quad (3)$$

where p is a pixel in the image; $H(p)$ is the adaptation factor at pixel p ; I_{CFA} is the intensity of the mosaic input image, normalized between [0, 1]; $*$ denotes the convolution operation; and G_H is a low-pass filter that models the transfer function of the horizontal cells. Here G_H is a two-dimensional Gaussian filter (Fig. 5) with spatial constant σ_H . For the images shown in this article, we used $\sigma_H=3$:

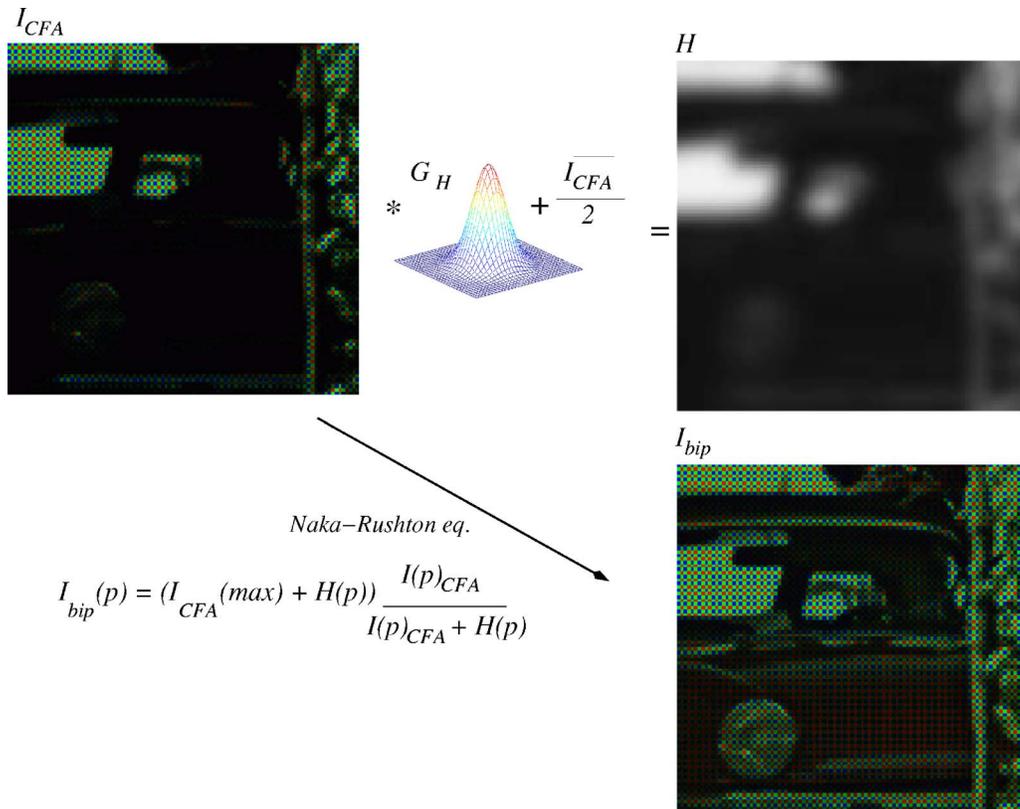


Fig. 5. (Color online) Simulation of the OPL adaptive nonlinear processing. The input signal is processed by the Naka-Rushton equation, whose adaptation factors are given by filtering the CFA image with a low-pass filter. The second nonlinearity that models the IPL layer works similarly.

$$G_H(x,y) = e^{-[(x^2+y^2)/2\sigma_H^2]}, \quad (4)$$

where $x \in [-4\sigma_H, 4\sigma_H]$ and $y \in [-4\sigma_H, 4\sigma_H]$.

The term $\overline{I_{CFA}}$ corresponds to the mean value of the CFA image pixel intensities. The factor (here $\frac{1}{2}$) induces different local effects and can be adjusted according to the image key. If we decrease the factor to a value closer to 0, the contrast in the shadows is enhanced, which might better render a low key image.

The input image I_{CFA} is then processed according to the Naka–Rushton equation (1) using the adaptation factors given by H . The responses of the bipolar cells network is computed with the following equation [Eq. (5)]. The parameters correspond to the mosaic and horizontal cell responses. A graphical representation is given in Fig. 5.

$$I_{bip}(p) = (I_{CFA}(\max) + H(p)) \frac{I_{CFA}(p)}{I_{CFA}(p) + H(p)}, \quad (5)$$

The term $(I_{CFA}(\max) + H(p))$ is a normalization factor that ensures that I_{bip} is again scaled in the range of [0, 1].

B. Second Nonlinearity

A second, similar nonlinear operation that models the behavior of the IPL is applied on the image I_{bip} to obtain the tone mapped image I_{ga} as

$$I_{ga}(p) = (I_{bip}(\max) + A(p)) \frac{I_{bip}(p)}{I_{bip}(p) + A(p)}, \quad (6)$$

where $A(p)$ simulates the output of the amacrine cells and I_{ga} models the output signal that would be transferred from the ganglion cells to the visual cortex. Similar to Eq. (3), A is a low-pass version of the image intensities at the bipolar cells level. It is computed by convolving the mosaic image I_{bip} with a Gaussian filter of spatial constant σ_A . We used $\sigma_A = 1.5$:

$$A(p) = I_{bip}(p) * G_A + \frac{\overline{I_{bip}}}{2}, \quad (7)$$

where G_A is given by

$$G_A(x,y) = e^{-[(x^2+y^2)/2\sigma_A^2]}, \quad (8)$$

and $x \in [-4\sigma_A, 4\sigma_A]$ and $y \in [-4\sigma_A, 4\sigma_A]$.

The resulting mosaic image I_{ga} has now been processed by a local tone mapping operator. Local contrast has been increased. The next step before displaying the result is to recover three chromatic components per spatial location. This can be performed by any demosaicing algorithm.

C. Demosaicing

We use the demosaicing algorithm described by Alleysson *et al.* [21], which first obtains the luminance image using a wide-band low-pass filter. Although some high frequencies are removed by this method [22], the filter is sufficiently accurate to well estimate the luminance. We chose a low-pass filter that removes even more high frequencies than the one presented in Alleysson *et al.*, as the two nonlinearities applied before already enhance the contours of the image. The implied difference of Gaussian (DOG) fil-

tering [12] results in a sharpening effect. In addition, removing high luminance frequencies also reduces noise.

We choose the luminance estimation filter to be F_{dem} :

$$F_{dem} = \frac{1}{256} \begin{bmatrix} 1 & 4 & 6 & 4 & 1 \\ 4 & 16 & 24 & 16 & 4 \\ 6 & 24 & 36 & 24 & 6 \\ 4 & 16 & 24 & 16 & 4 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix}. \quad (9)$$

Then

$$L(p) = I_{ga}(p) * F_{dem}, \quad (10)$$

where I_{ga} is the tone mapped CFA image and L represents the nonlinearly encoded luminance, which we call “lightness.” Note that in [21], L corresponds to the luminance, while here L is nonlinear and corresponds to perceived lightness. Nevertheless, the properties of the Fourier spectrum remain the same. We will use the term “lightness” to refer to L in the rest of the article.

The chrominance is then obtained by subtracting L from the mosaiced image I_{ga} :

$$C(p) = I_{ga}(p) - L(p), \quad (11)$$

where $C(p)$ is also a mosaic and contains the down-sampled chrominance. In $C(p)$, each pixel contains information only for one spectral band and can be separated into three downsampled chrominance channels using the modulation functions m_R , m_G , and m_B [see Eq. (12)]. This is illustrated in Fig. 6.

$$\begin{aligned} m_R(x,y) &= (1 + \cos(\pi x))(1 + \cos(\pi y))/4, \\ m_G(x,y) &= (1 - \cos(\pi x)\cos(\pi y))/2, \\ m_B(x,y) &= (1 - \cos(\pi x))(1 - \cos(\pi y))/4, \end{aligned} \quad (12)$$

where x,y is the coordinate of a pixel p in the image, with the upper-left pixel having coordinate 0,0. The chrominance channels are given by

$$C_1(x,y) = C(x,y) \cdot m_R(x,y),$$

$$C_2(x,y) = C(x,y) \cdot m_G(x,y),$$

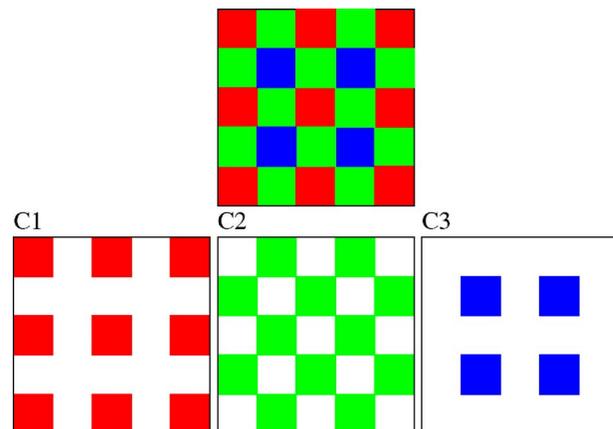


Fig. 6. (Color online) Chrominance channels are separated before interpolation.

$$C_3(x,y) = C(x,y) \cdot m_B(x,y). \quad (13)$$

In C_1, C_2, C_3 , the missing pixels (having a zero value) must be reconstructed to recover the full resolution image. This is done using a simple bilinear interpolation. Although more sophisticated methods exist, we deem it sufficient as the chrominances are isoluminant and do not contain high spatial frequencies [34].

After interpolation, the treated RGB image is obtained by adding the lightness and the chrominance channels together:

$$\begin{aligned} R(p) &= L(p) + C'_1(p), \\ G(p) &= L(p) + C'_2(p), \\ B(p) &= L(p) + C'_3(p), \end{aligned} \quad (14)$$

where $R(p)$, $G(p)$, $B(p)$ are the RGB channels of the image; L is the lightness [Eq. (10)]; and C'_1 , C'_2 , C'_3 are the interpolated chrominance channels.

4. RESULTS

We present results obtained with a Canon camera (Canon EOS 300D) and legacy images. In order to retrieve the RAW data, we used the free program DCRAW [35], which can handle RAW formats from nearly all cameras but does not apply color matricing or white balancing. Thus, to better illustrate the effect of the tone mapping algorithm alone, we present the results in black and white so that incomplete color rendering does not influence the visual results. Figure 2(d) shows a color example obtained from our algorithm.

To obtain simulated RAW images from legacy images, we inversed the original nonlinearity assuming a power function (gamma) [36] of 2.4 and recreated the mosaic according to the Bayer pattern.

The results for three scenes representing different dynamic ranges is shown in Fig. 7. The left and right images are legacy images. The image in the middle is a Canon RAW image. The results of our algorithm are compared with two center/surround local tone mapping algorithms: MSRCR (multiscale Retinex with color restoration) devel-



Fig. 7. Comparison of our algorithm with other tone mapping operators. Left column: Low-dynamic-range scene. Middle column: Medium-to high-dynamic-range scene. Right column: High-dynamic-range scene. First row: Global tone mapping with camera default setting. Second row: Images processed with MSRCR [30]. Third row: Images processed with the Retinex-based adaptive filter method [29]. Fourth row: Images processed with our proposed algorithm.

oped by Rahman *et al.* [30] and the adaptive filter method by Meylan and Süssstrunk [29]. The MSRCR image was obtained with the free version of the software “PhotoFlair” using the default settings [37] (which puts “demo” tags across the image). The globally corrected image (default camera settings) is also shown.

The advantage of our method is that it provides good-looking images regardless of the characteristics of the input image, while other methods are often restricted to a set of images having common features (dynamic range, key, and content). For example, MSRCR provides good tone mapping when the dynamic range is standard or slightly high, but it tends to generate artifacts when the input image has a very high dynamic range such as the one of Fig. 7, right-hand column, second row. The method is not able to retrieve all details in the center-right building, for example. The adaptive filter method [29] does not have these drawbacks but in general does not sufficiently increase local contrast in the light areas, which is visible in all images in the sky regions (Fig. 7, third row). Our method performs well for all three examples, the sky areas still have details, and the contrast in the dark areas is also enhanced.

In addition, another advantage of our method is that it is quite fast compared with other existing local tone mapping algorithms. First, the operation is performed on the CFA image, which divides the time of computation by three. Second, the fact that relatively small filters can be used for tone mapping (see Section 5) ensures that the algorithm has a reasonably low complexity.

5. DISCUSSION

We propose a tone mapping algorithm that is applied directly to the CFA image. It is inspired by a simple model of retinal processing that applies two nonlinearities on the spatially multiplexed chromatic signals. The nonlinearities are modeled with a Naka–Rushton function, where the adaptation parameter is an average of the local surround. It performs well in comparison with other local tone mapping algorithms.

Our interpretation of retinal processing is only partly supported by the literature on retina physiology. However, there are two processes supporting our hypothesis that can be found. First, there is a nonlinear process that occurs postreceptorally. Second, the role of horizontal cells that perform neighborhood connectivity is important for

the formation of the center/surround receptive fields present in the retina. As pointed out in Hood [7] (pp. 519–520), the formation of receptive fields is not yet completely understood. In particular, how horizontal cells modulate the cone responses is still under debate. We show here that using the horizontal cell responses to regulate the adaptive nonlinearity gives a good constraint on the signals and also prevents the appearance of artifacts. Finally, the hypothesis that the regulation in the IPL operates similarly to the one in the OPL is supported by studies that show a second nonlinearity in chromatic processing after the coding into opponent channels [18,19].

Section 4 compared the results of our algorithm with images obtained with other center/surround methods. We saw that our algorithm does not suffer from halos nor graying out and renders different scenes equally well. The reason why our method is more generally applicable is due to the fact that it is not based on the same general equation [Eq. (2)]. Indeed, with traditional methods, the local information is averaged and subtracted from the value of the treated pixel. Our algorithm also uses an average of the surrounding pixel values, given by H or A . However, it uses it as a variable in the Naka–Rushton equation (the adaptation factor), which is then applied to the treated pixel. If the treated pixel lies in a dark area, the adaptation factor is small and thus the output value range allocated to dark input values is large (Fig. 4). In a bright area, the adaptation factor is large and thus the mapping function between the input pixel value and the output pixel value is almost linear. This allows us to increase the local contrast in dark areas while still conserving local contrast in bright areas.

Another advantage of using such a technique is that the resulting image does not change much with different filter sizes. This makes our algorithm robust to varying parameters. In our implementation of the algorithm, we used $\sigma_H=3$ and $\sigma_A=1.5$. However, other values can be used without corrupting the results. Figure 8 shows an example of our method using different filter sizes, ($\sigma_H=1$; $\sigma_A=1$) for the left image and ($\sigma_H=3$; $\sigma_A=5$) for the right image. There is no tonal difference between the two resulting images. The slight discrepancy between the two images is due to the different sharpening effects induced by the change in filter size.

Our method aim was to achieve pleasing reproductions of images. This cannot be measured objectively. “Pleas-



Fig. 8. (Color online) Example of our method applied with different filter sizes. Left: Small filters ($\sigma_H=1$ and $\sigma_A=1$). Right: Large filter ($\sigma_H=3$ and $\sigma_A=5$).

ing” can mean different things to different people and is dependent not only on scene dynamic range and key but also on scene content. There are no objective criteria, and pleasantness should be evaluated using psychovisual experiments and human subjects. Previous evaluations of tone mapping algorithms, however, led to different conclusions depending not only on the scene content but also on the task [38,39]. Here, we provide a comparison with two other algorithms on three scenes. A few additional comparisons were published in a conference paper by Alleysson *et al.* [40]. We also made the code available online [41] so that figures and results are reproducible [42] for readers who wish to try our method on their own images.

6. CONCLUSION

We present a color image processing workflow that is based on a model of retinal processing. The principle of our workflow is to perform color rendering before color reconstruction (demaicing), which is coherent with the HVS. Our focus is on the tone mapping part of the general problem of color rendering. The integration of other rendering operations, such as white balancing and color matricing, is considered for future work.

Our proposed tone mapping algorithm is performed directly on the CFA image. It shares similarities with center/surround algorithms but is not subject to artifacts. The algorithm is fast compared with existing tone mapping methods and provides good results for all tested images.

ACKNOWLEDGMENTS

We would like to thank the anonymous reviewers for their pertinent comments that greatly improved the quality of the manuscript. This work was supported by the Swiss National Science Foundation under grant 21-101681.

REFERENCES

1. A. Roorda, A. B. Mehta, P. Lennie, and D. R. Williams, “Packing arrangement of the three cone classes in primate retina,” *Vision Res.* **41**, 1291–1306 (2001).
2. J. Holm, I. Tastl, L. Hanlon, and P. Hubel, “Color processing for digital photography,” in *Colour Engineering: Achieving Device Independent Colour*, P. Green and L. MacDonald eds. (Wiley, 2002), pp. 179–220.
3. C. R. Ingling and E. Martinez-Uriegas, “The spatiotemporal properties of the r-g x-cell channel,” *Vision Res.* **25**, 33–38 (1985).
4. R. L. De Valois and K. K. De Valois, *Spatial Vision*, Oxford Psychology Series 14, (Oxford U. Press, 1990).
5. N. V. S. Graham, *Visual Pattern Analysers*, Oxford Psychology Series 16 (Oxford U. Press, 1989).
6. K.-I. Naka and W. A. H. Rushton, “S-potentials from luminosity units in the retina of fish (Cyprinidae),” *J. Physiol. (London)* **185**, 587–599 (1966).
7. D. C. Hood, “Lower-level visual processing and models of light adaptation,” *Annu. Rev. Psychol.* **49**, 503–535 (1998).
8. G. J. Braun and M. D. Fairchild, “Image lightness rescaling using sigmoidal contrast enhancement functions,” *J. Electron. Imaging* **8**, 380–393 (1999).
9. J. Holm, “Photographic tone and colour reproduction goals,” in *Proceedings of CIE Expert Symposium’96 on Colour Standards for Image Technology*, (CIE, 1996), pp. 51–56.
10. R. Shapley and C. Enroth-Cugell, “Visual adaptation and retinal gain controls,” in *Progress in Retinal Research* (Pergamon, 1984), pp. 263–346.
11. R. M. Haralick and L. G. Shapiro, *Computer and Robot Vision*, 1st ed. (Addison-Wesley, 1993).
12. W. K. Pratt, *Digital Image Processing* (Wiley, 1991).
13. A. Roorda and D. R. Williams, “The arrangement of the three cone classes in the living human eye,” *Nature* **397**, 520–522 (1999).
14. B. E. Bayer, “Color imaging array,” U.S. patent 3,971,065 (March 3, 1976).
15. J. H. Van Hateren, “Encoding of high dynamic range video with a model of human cones,” *ACM Trans. Graphics* **25**, 1380–1399 (2006).
16. See <http://www.webvision.med.utah.edu/>.
17. M. Kamermans and H. Spekrijse, “The feedback pathway from horizontal cells to cones. A mini review with a look ahead,” *Vision Res.* **39**, 2449–2468 (1999).
18. M. A. Webster and J. Mollon, “Changes in colour appearance following post-receptoral adaptation,” *Nature (London)* **349**, 235–238 (1991).
19. T. Yeh, J. Pokorny, and V. C. Smith, “Chromatic discrimination with variation in chromaticity and luminance: data and theory,” *Vision Res.* **33**, 1835–1845 (1993).
20. H. Spitzer and S. Semo, “Color constancy: a biological model and its application for still and video images,” *Pattern Recogn.* **35**, 1645–1659 (2002).
21. D. Alleysson, S. Süssstrunk, and J. Herault, “Linear demosaicing inspired by the human visual system,” *IEEE Trans. Image Process.* **14**, 439–449 (2005).
22. E. Dubois, “Frequency-domain methods for demosaicing of Bayer-sampled color images,” *IEEE Signal Process. Lett.* **12**, 847–850 (2005).
23. N. Lian, L. Chang, and Y. Tan, “Improved color filter array demosaicing by accurate luminance estimation,” in *Proceedings of IEEE Conference on Image Processing (IEEE, 2005)*, pp. I-41–44.
24. E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, “Photographic tone reproduction for digital images,” in *Proceedings of ACM SIGGRAPH 2002, Annual Conference on Computer Graphics* (ACM, 2002), pp. 267–276.
25. G. Ward, H. Rushmeier, and C. Piatko, “A visibility matching tone reproduction operator for high dynamic range scenes,” *IEEE Trans. Vis. Comput. Graph.* **3**, 291–306 (1997).
26. K. Devlin, “A review of tone reproduction techniques,” Technical Rep. CSTR-02-005 (Department of Computer Science, University of Bristol, 2002).
27. E. Reinhard, G. Ward, S. Pattanaik, and P. Debevec, *High Dynamic Range Imaging. Acquisition, Display, and Image-Based Lighting*, (Morgan Kaufmann, 2005).
28. M. Ashikhmin, “A tone mapping algorithm for high contrast images,” in *Proceedings of Eurographics Workshop on Rendering Techniques* (Eurographics Association, 2002), pp. 145–155.
29. L. Meylan and S. Süssstrunk, “High dynamic range image rendering with a Retinex-based adaptive filter,” *IEEE Trans. Image Process.* **15**, 2820–2830 (2006).
30. Z.-U. Rahman, D. J. Jobson, and G. A. Woodell, “Retinex processing for automatic image enhancement,” *J. Electron. Imaging* **13**, 100–110 (2004).
31. R. Fattal, D. Lischinski, and M. Werman, “Gradient domain high dynamic range compression,” in *Proceedings of ACM SIGGRAPH 2002, Annual Conference on Computer Graphics* (ACM, 2002), pp. 249–256.
32. F. Durand and J. Dorsey, “Fast bilateral filtering for the display of high-dynamic-range images,” in *Proceedings of ACM SIGGRAPH 2002, Annual Conference on Computer Graphics* (ACM, 2002), pp. 257–266.
33. K. Barnard and B. Funt, “Investigations into multi-scale Retinex,” in *Colour Imaging: Vision and Technology*, (Wiley, 1999), pp. 9–17.
34. K. T. Mullen, “The contrast sensitivity of human colour vision to red/green and blue/yellow chromatic gratings,” *J. Physiol. (London)* **359**, 381–400 (1985).
35. D. Coffin, <http://cybercom.net/~dcoffin/dcrawl/>.

36. IEC, "Multimedia systems and equipment—colour measurement and management—Part2-1:colour management—default RGB colour space—sRGB," IEC61966-2-1 (IEC, 1999).
37. Truview Imaging Company, <http://truview.com>.
38. P. Ledda, A. Chalmers, T. Troscianko, and H. Seetzen, "Evaluation of tone mapping operators using a high dynamic range display," in *Proceedings of ACM SIGGRAPH 2005, Annual Conference on Computer Graphics* (ACM, 2005), pp. 640–648.
39. J. Kuang, H. Yamaguchi, G. M. Johnson, and M. D. Fairchild, "Testing HDR image rendering algorithms," in *Proceedings of IS&T/SID Twelfth Color Imaging Conference: Color Science, Systems, and Application*, (IS&T, 2004), pp. 315–320.
40. D. Alleysson, L. Meylan and S. Ssstrunk, "HDR CFA image rendering," in *Proceedings of EURASIP 14th European Signal Processing Conference* (European Association for Signal Processing, 2006).
41. Supplementary material available at http://ivrg.epfl.ch/supplementary_material/index.html.
42. M. Schwab, M. Karrenbach, and J. Claerbout, "Making scientific computations reproducible," *Comput. Sci. Eng.* **2**, 61–67 (2000).