

A Model of Retinal Local Adaptation for the Tone Mapping of Color Filter Array Images

Laurence Meylan, David Alleysson, Sabine Süsstrunk

▶ **To cite this version:**

Laurence Meylan, David Alleysson, Sabine Süsstrunk. A Model of Retinal Local Adaptation for the Tone Mapping of Color Filter Array Images. Journal of the Optical Society of America. A Optics, Image Science, and Vision, 2007, 24 (9), pp.2807-2816. 10.1364/josaa.24.002807 . hal-00204914

HAL Id: hal-00204914
https://hal.science/hal-00204914

Submitted on 15 Jan 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Model of Retinal Local Adaptation for the Tone Mapping of Color Filter Array Images

Laurence Meylan

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

David Alleysson

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

Sabine Süsstrunk

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

laurence.meylan@a3.epfl.ch; david.alleysson@upmf-grenoble.fr; sabine.susstrunk@epfl.ch

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, analogue to the human visual system that applies a non-linearity on the chromatic responses captured by the cone mosaic. This reduces the number of necessary operations by a factor three. 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 using it 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.

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 pixel (Fig. 1, left). 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 work-flows,[1] 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 work-flow shown in Fig. 2 (a), we propose a solution where the demosaicing is the last step of the color processing work-flow. Color rendering operations are thus performed directly on the CFA image (Fig. 2, b). In this article, we only consider 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 work-flow is that it is more analogous to the retinal processing of the human visual system (HVS),[2–4] as discussed in Section 2. Another motivation is that applying the tone mapping directly on 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 to other existing 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 non-linear adaptation that occurs in the retina, which efficiently improves local contrasts while conserving good global appearance.[5,6] Fig. 2 (c,d) shows an example of applying our method on 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[7,8] (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, they are well rendered without requiring an additional sharpening operation.

We applied our algorithm on 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 while not creating

artifacts.

This article is structured as follows. Section 2 provides background knowledge on tone mapping and the model of retinal adaptation that we base our method on. Section 3 presents the algorithm. Section 4 shows the results obtained by our proposed work-flow, and Section 5 discusses the differences of our algorithm compared to 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 non-linearity that applies on that mosaic. We concentrate on one specific non-linear processing model proposed by Naka and Rushton[5,9] 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, and specifically the center/surround family of local tone mapping algorithms is also reviewed, as our method bears some similarity to the latter.

2.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.[10,11] Another equivalence exists between the spatio-chromatic sampling of the cone mosaic and the sampling of color in single chip sensor such as given by the Bayer CFA (Fig. 1).[12,13]

Our proposed work-flow (Fig. 2, b) exploits another analogy with human vision, namely between the tone mapping operations in the image processing work-flow and the non-linear 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.[14] We aim to identify, and simplify, which processing acts on the retinal signal in order to develop algorithms suitable for in-camera processing. We focus on the non-linearities applied to the mosaic of chromatic responses captured by the cones.

One role of tone mapping is to non-linearly process the captured image to mimic the retina's non-linear adaptation and render the image as if the HVS had processed it. In traditional work-flows, this non-linear encoding is usually applied to the RGB color image, thus after the color mosaic captured by the CFA sensor is demosaiced. For the HVS, the non-linear 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, 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[2] (see Fig. 3). We thus propose a new image processing work-flow where the non-linear encoding (tone mapping) is directly performed on the mosaic image provided by the Bayer CFA pattern.

Fig. 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[15]). 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 non-linearity on the input signal. These two layers are composed of the cones, the horizontal and amacrine cells, which provide the horizontal connectivity, and of 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.[16] 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 non-linear processing is applied. We assume that the amacrine cells also provide a feed-back to modulate the adaptive non-linearity of the ganglion cells. This second non-linearity has been found to provide psychophysical[17,18] and physiological[9] evidence for an adaptation mechanism to contrast rather than to intensity. Moreover, it has been suggested that this non-linearity is postreceptoral and applies on color opponent representation.[6,18] We assume here that it originates in the interaction between bipolar, amacrine, and ganglion cells.

Our tone mapping algorithm also applies two non-linear processings on the CFA image in imitation of the IPL and OPL functionalities. Both non-linear operations are based on Naka and Rushton,[5,9] who developed a model for the photoreceptor non-linearities and adaptation to incoming light. Spitzer et al.[19] also proposed a biological model for color contrast, which used similar adaptation mechanisms. The non-linear mosaic image is then demosaiced to reconstruct the RGB tone-mapped image.

2.B. Adaptive Non-Linearity

Our model of the OPL and IPL non-linearities takes inspiration from the Naka-Rushton equation[5,9]

$$Y = \frac{X}{X + X0}, \tag{1}$$

where X represents the input light intensity, $X0$ is the adaptation factor, and Y is the adapted signal. In the original formulation,[5] the adaptation factor ($X0$) is determined by the average light reaching the entire field of view. In our method, $X0$ varies for each pixel.

It is a local variable given by the average light intensity in the neighborhood of one pixel. Fig. 4 illustrates the Naka-Rushton function for different values of $X0$. If $X0$ is small, the cell output has increased sensitivity. If $X0$ is large, there is not much change in sensitivity.

In our model, the Naka-Rushton equation is used to calculate the non-linearities of both the OPL and IPL. $X0$ 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 non-linearity is assumed to be due only to the dynamics of the photoreceptors themselves.[14] 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 feedforward on bipolar cells.[16] Thus, our assumption is that the mechanism by which horizontal cells modify cone responses is due to a regulation of the cone's non-linear adaptation factor, based on the response of the horizontal cells network at the cone location.

2.C. Properties of a CFA Image

The two non-linearities described above are directly applied on the CFA image. In our implementation, the CFA image is obtained using a Bayer pattern[13] in front of the camera sensor, which results in a spatio-chromatic sampling of the scene. This mosaic image has certain properties that allows treating the luminance and the chrominance of the image separately.

Alleysson et al.[20] 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 down-sampled and encoded with opponent colors. It follows that a wide-band low-pass filter can be used to recover the luminance and a high-pass or band-pass filter can recover the down-sampled chrominance. Choosing the appropriate filters allows implementing an efficient demosaicing algorithm. Their method was refined by Dubois[21] and Lian et al.[22] who propose a more accurate estimation of the luminance.

In Section 3.C, we will apply the Alleysson et al. method for demosaicing. In Section 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 non-linearities only on the luminance signals, not on any chromatic components.

2.D. Tone Mapping

Tone mapping is the operation in the image processing work-flow 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 "s-shape." More sophisticated global tone mapping methods vary the function parameters depending on global characteristics of the image.[7,8,23,24] The key of the image can be used to determine the exponent of the gamma function.[23] In Braun and Fairchild[7] and in Holm,[8] a s-shaped function is defined by the image statistics, such as the mean and the variance of the intensity. In Ward et al.,[24] 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 in different classes sharing the same common features (see Delvin[25] and Reinhard *et al.* [26] 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.[23,27–29] 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 Section 2.E. *Gradient-based* methods[30] 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 to integrate the gradient to recover the treated image. *Frequency-based* methods[31] separate the low and high frequency bands of the image. The low-frequency band is assumed to approximatively correspond 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 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 to increase the local contrast and keep detail visibility.

2.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.

$$I'(p) = log(I(p)) - log(I(p) * G), \tag{2}$$

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 literatures[28,29,32] and solutions to overcome them were developed. Rahman et al.[29] 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[28] 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 re-computed 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, artifacts-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 non-linear function rather than as a fixed factor subtracted

from the input pixel.

3. A Local Tone Mapping Algorithm for CFA Images

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

3.A. The First Non-Linearity

The first non-linear operation simulates the adaptive non-linearity 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.

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

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. G_H is a two-dimensional Gaussian filter (Fig. 5) with spatial constant σ_H. For the images shown in this article, we used $\sigma_H = 3$.

$$G_H(x, y) = e^{-\frac{x^2+y^2}{2\sigma_H^2}}, \tag{4}$$

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 Equ. (1) using the adaptation factors given by H. The responses of the bipolar cells network is computed with the following equation (Equ. 5). The parameters correspond to the mosaic and horizontal cell responses. A graphical representation is given in Fig. 5.

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

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

3.B. The Second Non-Linearity

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

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

$A(p)$ simulates the output of the amacrine cells. I_{ga} models the output signal that would be transfered from the ganglion cells to the visual cortex. Similarly to Equ. (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 $\sigma_A = 1.5$.

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

where G_A is given by

$$G_A(x, y) = e^{-\frac{x^2+y^2}{2\sigma_A^2}}, \tag{8}$$

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.

3.C. Demosaicing

We use the demosaicing algorithm described by Alleysson et al.,[20] which first obtains the luminance image using a wide-band low-pass filter. Although some high frequencies are removed by this method,[21] 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 non-linearities applied before already enhance the contours of the image. The implied Difference of Gaussian (DOG) filtering[11] results in a sharpening effect. In addition, removing high luminance frequencies also reduces noise.

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

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

Then

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

where I_{ga} is the tone-mapped CFA image and L represents the non-linearly encoded luminance, which we call "lightness". Note that in,[20] L corresponds to the luminance while here L is non-linear 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}.

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

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

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

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:

$$\begin{aligned} C_1(x,y) &= C(x,y) \cdot m_R(x,y) \\ C_2(x,y) &= C(x,y) \cdot m_G(x,y) \\ C_3(x,y) &= C(x,y) \cdot m_B(x,y) \end{aligned} \tag{13}$$

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.[33]

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

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

where $R(p)$, $G(p)$, $B(p)$ are the RGB channels of the image, L is the lightness (Equ. 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,[34] 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. Fig. 2 (d) shows a color example obtained from our algorithm.

To obtain simulated RAW images from legacy images, we inversed the original non-linearity assuming a power function (gamma)[35] 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 to two center/surround local tone mapping algorithms: MSRCR (Multi-Scale Retinex with Color Restoration) developed in Rahman et al.,[29] and the adaptive filter method by Meylan and Süsstrunk.[28] The MSRCR image was obtained with the free version of the software "PhotoFlair" using the default settings,[36] (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, 2nd row. The method is not able to retrieve all details in the center-right building, for example. The adaptive filter method[28] 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, 3rd 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 to 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 directly applied on the CFA image. It its inspired by a simple model of retinal processing that applies two non-linearities on the spatially multiplexed chromatic signals. The non-linearities are modeled with a Naka-Rushton function, where the adaptation parameter is an average of the local surround. It performs well in comparison to 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 non-linear process that occurs post-receptorally. 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[6] (pages 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 non-linearity gives a good constraint on the signals and also prevents the apparition 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 non-linearity in chromatic processing after the coding into opponent channels.[17,18]

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 (Equ. 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 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. Fig. 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 aims to achieve pleasing reproductions of images. This can not be measured objectively. "Pleasing" can mean different things to different people, and is not only dependent 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.[37,38] Here, we provide a comparison with two other algorithms on three scenes. We put the code available on-line[39] so that figures and results are reproducible[40] and for readers who wish to try our method on their own images.

6. Conclusion

We present a color image processing work-flow that is based on a model of retinal processing. The principle of our work-flow is to perform color rendering before color reconstruction (demosaicing), 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 to existing tone mapping methods and provides good results for all tested images.

References

1. J. Holm, I. Tastl, L. Hanlon, and P. Hubel. *Colour Engineering: Achieving Device Independent Colour; ch. Color processing for digital photography*, pages 179–220. P. Green and L. MacdDonald ed., John Wiley and Sons, 2002.
2. C. R. Ingling and E. Martinez-Uriegas. "The spatiotemporal properties of the r-g x-cell channel." *Vision Research*, **25**(1):33–38, 1985.
3. R. L. De Valois and K. K. De Valois. *Spatial Vision*. Oxford Psychology Series, 14, 1990.
4. N. V. S. Graham. *Visual Pattern Analysers*. Oxford Psychology Series, 16, 1989.
5. K.-I. Naka and W. A. H Rushton. "S-potentials from luminosity units in the retina of fish (cyprinidae)." *Journal of Physiology*, **185**(3):587–599, 1966.
6. D. C. Hood. "Lower-level visual processing and models of light adaptation." *Annual Review of Psychology*, **49**:503–535, 1998.
7. G. J. Braun and M. D. Fairchild. "Image lightness rescaling using sigmoidal contrast enhancement functions." *Journal of Electronic Imaging*, **8**(4):380–393, October 1999.
8. J. Holm. "Photographic tone and colour reproduction goals." In *Proc. CIE Expert Symposium'96 on Colour Standards for Image Technology*, pages 51–56, 1996.

9. R. Shapley and C. Enroth-Cugell. *Visual adaptation and retinal gain controls. Progress in retinal research*, pages 263–346. Pergamon Press, 1984.
10. R. M. Haralick and L. G. Shapiro. Computer and Robot Vision. 1st edition, Addison-Wesley Longman Publishing Co., 1993.
11. W .K. Pratt. *Digital Image Processing.* Wiley, New York, 1991.
12. A. Roorda and D. R. Williams. "The arrangement of the three cone classes in the living human eye." *Nature*, **397**(11), 1999.
13. B. E. Bayer. "Color imaging array." US patent #3,971,065 Eastman Kodak Company, 1976.
14. J. H. Van Hateren. "Encoding of high dynamic range video with a model of human cones." *ACM Transaction on Graphics*, **25**(4):1380-1399, October 2006.
15. http://webvision.med.utah.edu/.
16. M. Kamermans and H. Spekreijse. "The feedback pathway from horizontal cells to cones. A mini review with a look ahead." *Vision Research*, **39**(15):2449–2468, July 1999.
17. M. A. Webster and J. Mollon. "Changes in colour appearance following post-receptoral adaptation." *Nature*, **349**: 235 - 238, January 1991.
18. T. Yeh, J. Pokorny, and V.C. Smith. "Chromatic discrimination with variation in chromaticity and luminance: Data and theory." *Vision Research*, **33**(13):1835-1845, September 1993.
19. H. Spitzer and S. Semo. "Color constancy: a biological model and its application for still and video images." *Pattern Recognition*, **35**(8):1645–1659, 2002.
20. D. Alleysson, S. Süsstrunk, and J. Herault. "Linear demosaicing inspired by the human visual system." *IEEE Transactions on Image Processing*, **14**(4):439–449, April 2005.
21. E. Dubois. "Frequency-domain methods for demosaicking of bayer-sampled color images." *IEEE Signal Processing Letters*, **12**(12):847–850, December 2005.
22. N. Lian, L. Chang, and Y. Tan. "Improved color filter array demosaicing by accurate luminance estimation." In *Proc. IEEE International Conference on Image Processing, ICIP 2005*, **1**:I-41-4, 2005.
23. E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda. "Photographic tone reproduction for digital images." In *Proc. ACM SIGGRAPH 2002, Annual Conference on Computer Graphics*, pages 267–276, 2002.
24. G. Ward, H. Rushmeier, and C. Piatko. "A visibility matching tone reproduction operator for high dynamic range scenes." *IEEE Transactions on Visualization and Computer Graphics*, **3**(4):291–306, October-December 1997.
25. K. Devlin. A review of tone reproduction techniques. Technical Report CSTR-02-005, Department of Computer Science, University of Bristol, November 2002.
26. E. Reinhard, G. Ward, S. Pattanaik, and P. Debevec. *High Dynamic Range Imaging.*

Acquisition, Display, and Image-Based Lighting. Morgan Kaufmann Publishers, 2005.

27. M. Ashikhmin. "A tone mapping algorithm for high contrast images." In *Proc. Thirteenth Eurographics Workshop on Rendering*, pages 145–155, 2002.
28. L. Meylan and S. Süsstrunk. "High dynamic range image rendering with a Retinex-based adaptive filter." *IEEE Transactions on Image Processing*, **15**(9):2820–2830, September 2006.
29. Z.-U. Rahman, D. J. Jobson, and G. A. Woodell. "Retinex processing for automatic image enhancement." *Journal of Electronic Imaging*, **13**(1):100–110, January 2004.
30. R. Fattal, D. Lischinski, and M. Werman. "Gradient domain high dynamic range compression." In *Proc. ACM SIGGRAPH 2002, Annual Conference on Computer Graphics*, pages 249–256, 2002.
31. F. Durand and J. Dorsey. "Fast bilateral filtering for the display of high-dynamic-range images." In *Proc. ACM SIGGRAPH 2002, Annual Conference on Computer Graphics*, pages 257–266, 2002.
32. K. Barnard and B. Funt. *Colour imaging: vision and technology; ch. Investigations into multi-scale Retinex.*, pages 9–17. John Wiley and Sons, 1999.
33. K. T. Mullen, "The contrast sensitivity of human colour vision to red/green and blue/yellow chromatic gratings.". *Journal of Physiology*, **359**:381–400, 1985.
34. D. Coffin. http://cybercom.net/d̃coffin/dcraw/.
35. IEC 61966-2-1. Multimedia systems and equipment - Colour measurement and management - Part2-1:Colour management - Default RGB colour space - sRGB, 1999.
36. Truview imaging company. http://trueview.com.
37. P. Ledda, A. Chalmers, T. Troscianko, and H. Seetzen. "Evaluation of tone mapping operators using a high dynamic range display." In *Proc. ACM SIGGRAPH 2005, Annual Conference on Computer Graphics*, pages 640–648, 2005.
38. J. Kuang, H. Yamaguchi, G. M. Johnson, and M. D. Fairchild. "Testing HDR image rendering algorithms." In *Proc. IS&T/SID Twelfth Color Imaging Conference: Color Science, Systems, and Application*, pages 315–320, 2004.
39. Supplementary material. http://ivrg.epfl.ch/supplementary_material/index.html, 2007.
40. M. Schwab, M. Karrenbach, and J. Claerbout. "Making scientific computations reproducible." *Computing in Science & Engineering*, **2**(6):61–67, November 2000.

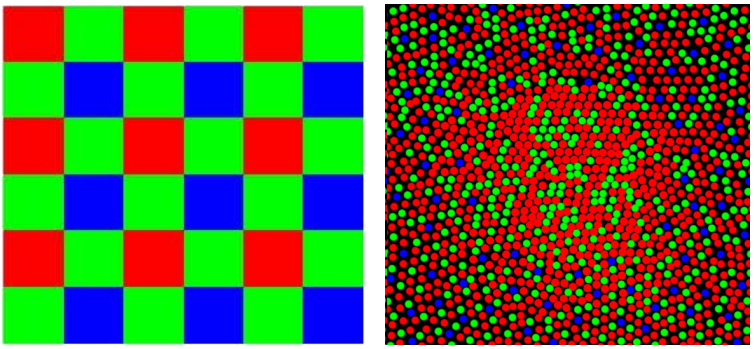
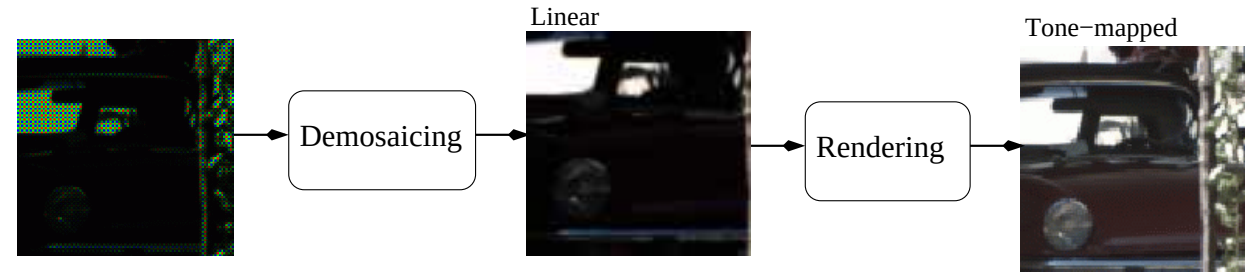
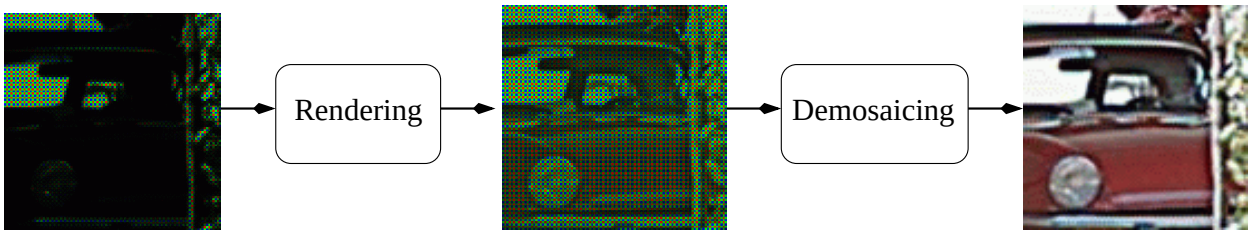


Fig. 1. (color online) Bayer CFA (left) and the spatio-chromatic sampling of the cone mosaic (right) [Inspired from Roorda et al. Vision Research, 2001].



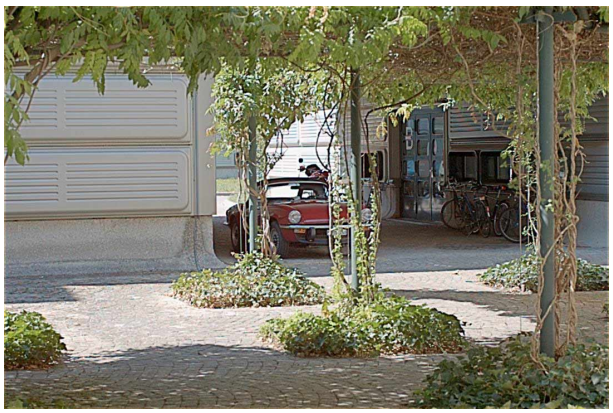
a) Traditional image processing workflow



b) Our proposed workflow



c) Global correction (gamma)



d) Our proposed method

Fig. 2. (color online) **Top (a):** Traditional image processing work-flow. **Center (b):** Our proposed work-flow. **Bottom left (c):** Image rendered with a global tone mapping operator (gamma). **Bottom right (d):** Image rendered according to our method.

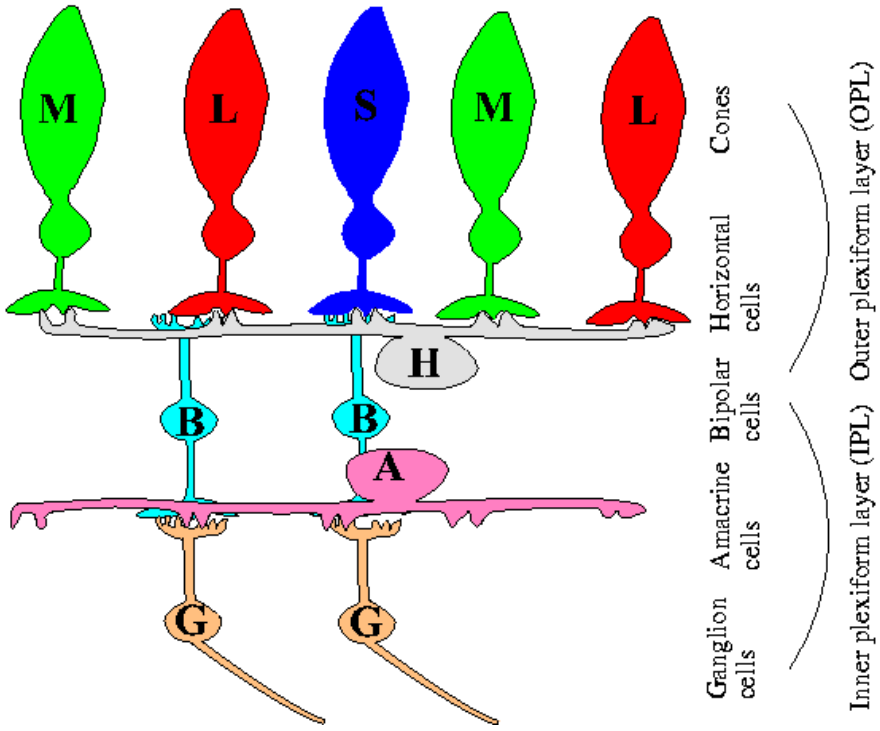


Fig. 3. (color online) Simplified model of the retina.

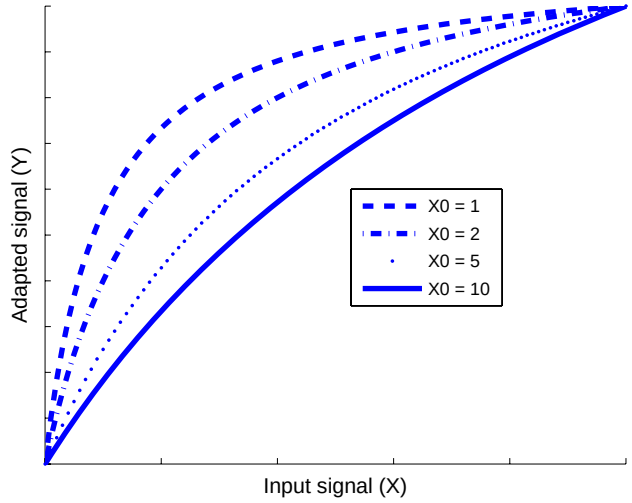


Fig. 4. Naka-Rushton function with different adaptation factors $X0$.

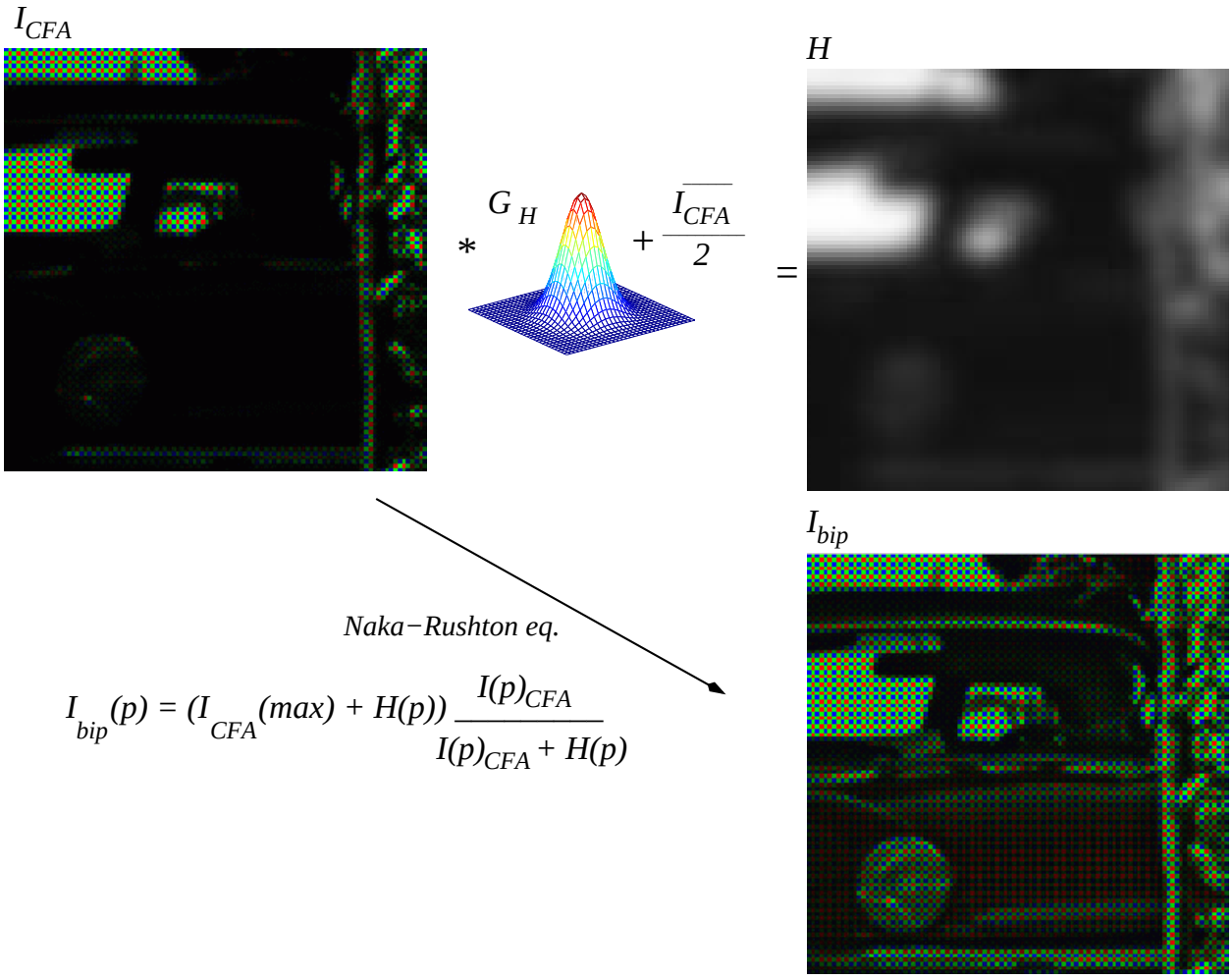


Fig. 5. (color online) Simulation of the OPL adaptive non-linear processing. The input signal is processed by the Naka-Rushton equation whose adaptation factors are given by filtering the CFA image with a low-pass filter. The second non-linearity that models the IPL layer works similarly.

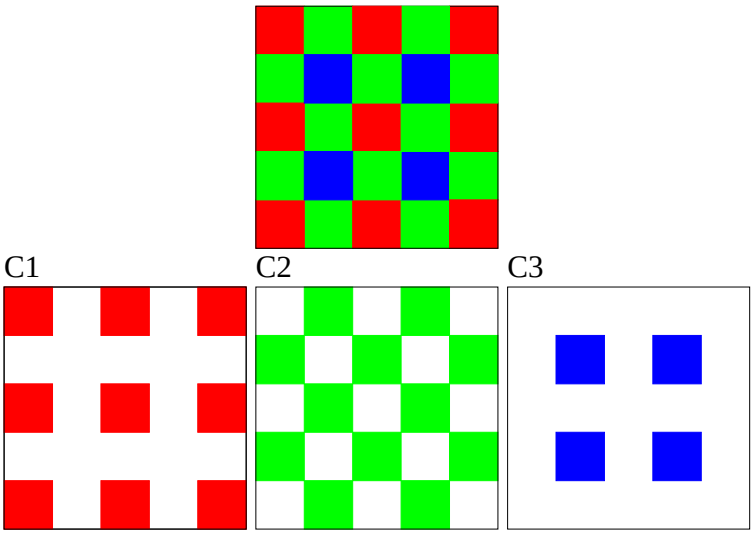


Fig. 6. (color online) The chrominance channels are separated before interpolation.



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.[29] **Third row:** Images processed with the Retinex-based adaptive filter method.[28] **Fourth row:** Images processed with our proposed algorithm.



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 $\sigma_A = 5$).