HDR CFA IMAGE RENDERING

David Alleysson,[1] *Laurence Meylan,*[2] *and Sabine Süsstrunk*[2]

[1]Laboratory of Psychology and Neurocognition
CNRS UMR 5105
University Pierre Mendes-France Grenoble, France
email: david.alleysson@upmf-grenoble.fr
[2]School of Computer and Communication Sciences (IC)
Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland
email: laurence.meylan, sabine.susstrunk@epfl.ch

ABSTRACT

We propose a method for high dynamic range (HDR) mapping that is directly applied on the color filter array (CFA) image instead of the already demosaiced image. This rendering is closer to retinal processing where an image is acquired by a mosaic of cones and where adaptive non-linear functions apply before interpolation. Thus, in our framework, demosaicing is the final step of the rendering. Our method, inspired by retinal sampling and adaptive processing is very simple, fast because only one third of operations are needed, and gives good result as shown by experiments.

1. INTRODUCTION

Most of todays digital cameras use a single sensor coupled with a Color Filter Array (CFA) for sensing colors. Furthermore, the acquired signal is quantized to more than 8 bits to allow for a noise reduced non-linear image rendering process. Often, digital camera image processing starts with demosaicing, a process for color interpolation, which is followed by color correction and tone mapping. In this paper, we investigate rendering of the images before demosaicing. Our method is based on a model of the human retina, where colors are also acquired by a mosaic of cone photoreceptors and where adaptive non-linear processes occur before interpolation.

After acquiring an image with a digital camera, some processing is needed to render the image “pleasing” to the observer on a given output device. Such processing includes white-balancing, color correction, and tone mapping. We are focusing on the latter in this paper.

If the dynamic range of the output device is similar to the dynamic range of the scene (or the focal plane irradiance), then a global tone mapping is usually sufficient to render the scene luminances. Global tone mapping operators compress the dynamic range non-linearly, using for example a logarithmic, gamma, or sigmoidal function [6]. However, if the dynamic range of the scene by far exceeds the dynamic range of the output device, applying only a global operator tends to compress the tonal range too much, causing a loss of contrast that results in a loss of detail visibility (see Figure 4 b), which we often interpret as regions of under- or overexposure.

The human visual system (HVS), on the other hand, is quite able to process HDR scenes without loss of contrast, as it can adapt to several orders of magnitudes of light intensity. On a basic level, the eye contains two sensor systems, cones and rods, each with functioning ranges for daylight and nocturnal vision. However, the HVS is also able to fully interpret the wide range of luminance levels that occur in daylight conditions. Compared to digital camera images with only global tone compression, where under- and overexposed regions are common, the human visual system always give us good detail discrimination.

It is meaningful to understand which kind of processing the HVS operates on the acquired light irradiances and to convert it into algorithms for digital images. Thus, most existing HDR algorithms [17, 19, 21] are based on HVS models, such as Retinex [12, 13], which simulate the local adaptation that occurs while the eye scans a scene. Recently proposed methods not only use a HVS model, but additionally explicitly mimic its functionality by implementing neural processes. In [14], Ledda et al. propose, for example, to use a model of cone and rod photoreceptors to simulate a local eye adaptation on HDR images. Reinhard and Delvin [18] propose the use of a model of cone physiology for dynamic range reduction in daylight conditions. These two approaches allow refining the modeling of visual processes rather than adapting parameters, such as is done in color appearance models [8].

For our HDR rendering framework, we consider a more extend model of the retina by taking into account several layers of neurons. Also, we use the similarity between a CFA and cone image, namely the sampling of just one single chromatic value per spatial position. Note that in the retina, this cone image or mosaic is known to be random [20], while the camera mosaics are generally regular. Despite this difference, we think that applying the HDR rendering on the mosaiced image better resembles the visual system than applying demosaicing first. In terms of computation, working on the mosaiced image reduces computational complexity because there are three times less pixels to process.

2. MODEL OF RETINAL PROCESSING

It should be noted that we still know very little about the processing of visual information by the HVS, and what we do know concerns mostly the retinal processes. The retina can be studied in isolation by comparing its output to a calibrated input. Even if there are many physiological studies, modeling the HVS behavior is always subject to interpretation. As an illustration, the adaptive and non-linear response of photoreceptors has been measured with flash illumination in isolated photoreceptor [9]. But we know that photoreceptors are coupled with each other [5] and are part of a synaptic triad where photoreceptor, horizontal, and bipolar cells form a dense group. We can thus question the plausibility of the

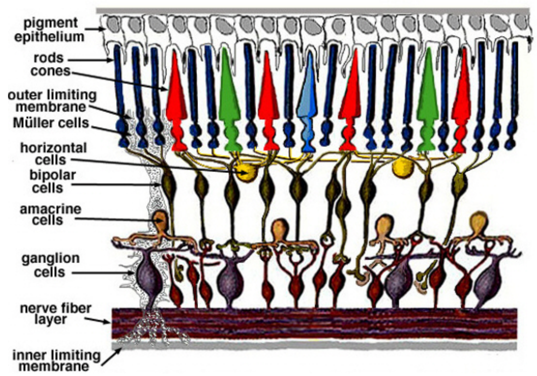

Figure 1: Schema of the human retina (From webvision [11])

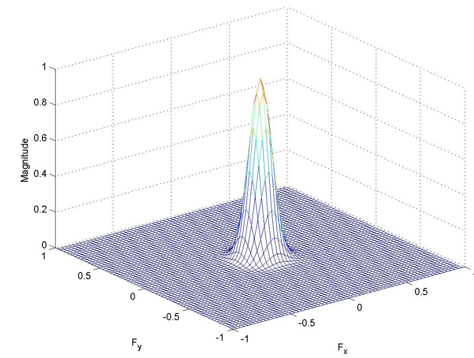

Figure 2: Impulse response of the horizontal cells filter.

model of isolated flash responses in this context. Nevertheless, physiological HVS models offer an interesting illustration of the visual processing from which we can get inspiration for digital imaging algorithms.

The retina is composed of two main layers (Figure 1), the inner plexiform layer (IPL) which is the location of the synaptic triad of cone, horizontal, and bipolar cells, and the outer plexiform layer (OPL) where bipolar, amacrine, and ganglion cells communicate. In these layers, horizontal and amacrine cells have a role of horizontal connectivity, where bipolar cells transmit signals from the IPL to OPL. Cones sample light and ganglion cells transmit information to the cortex. The definitive role of this cell network is not known and remains controversial.

Our HDR algorithm is based on a simple model of retinal processing, consisting of a mosaic of chromatic samples on which we apply two non-linear adaptive processes representing the IPL and the OPL. In the final step we apply demosaicing to render the full color image.

2.1 Chromatic mosaic image formation

In the retina, there are three kinds of cones active in daylight vision, called L, M and S for Long, Middle and Short wavelength spectral sensitivity, respectively. These cones form a random mosaic and as a consequence there is only a single chromatic response at each spatial location of the retina. Most digital cameras similarly sample only a single chromatic sensitivity per spatial location through a CFA. Thus, there is an analogy between the retina and digital camera sampling of chromatic information.

Moreover, there is no physiological evidence that the image is reconstructed as a regular image in the retina, i.e. an image with three chromatic samples per spatial location. On the contrary, it seems that at least in the parvocellular channel the chromatic and achromatic information remains multiplexed [10]. Thus, it seems reasonable to assume that the visual processing in the retina operates directly on the mosaic of cone responses.

In digital cameras, the mosaicing process according to the Bayer CFA is comparable to a frequency modulation of the chromatic signals. This modulation has the property to modulate the chrominance in the border of the Fourier spectrum and leaves the luminance of the image (located in the middle of the spectrum) unchanged [2]. Thus, we can design filters that apply independently on luminance and chrominance of the CFA image following the kind of filter (low or high pass) we design. This is similar to some HDR rendering algorithms, where the image is first transformed to a luminance chrominance representation, and the local tone mapping is only applied to the luminance [15].

2.2 Horizontal cell processing

We suppose that the role of horizontal cells is to estimate a spatio-temporal low pass filter of the CFA image. Since the filter is low pass, it applies only on the luminance information of the CFA. This is supported by their non-opponent response to visual stimuli [4]. Also, we propose that this filter has a small cut-off frequency, according to the size of the receptive field structure of these cells [16]. We thus use a FIR filter of size 33x33 given by the transfer function shown in Figure 2.

2.3 Adaptive non-linearity

We use an adaptive non-linearity that allows adapting the level of the signal with a non linear mapping. As already proposed by others [14, 18], this adaptive non-linearity can be implemented with a photoreceptor model given by the Naka-Rushton equation.

$$y = k\frac{x}{x+x_0} \tag{1}$$

where x is the input light intensity, and y is the output light intensity, x_0 is the adaptation factor and k is a gain factor for a digital value range between $[0,M]$, $M = 2^{16}$ for 16 bits images. We want the function of Equation 1 to return a value M for an input of value of M. Thus, k acts as a range normalization factor: $k = M + x_0$.

The parameter x_0 can be chosen either as a local or global parameter. The local behavior is given by the horizontal cells, which are known to have a feedback process on cones [3]. We suppose that this feedback modulates the adaptation parameter of the cones. The global factor is given by the mean of pixel intensities over the whole CFA image.

$$x_0 = F_h * x + \bar{x}/2 \tag{2}$$

where $F_h * x$ is the signal corresponding to the filtering of the CFA pixel intensity x by the transfer function F_h of the horizontal cell layer. This factor is local because its level depends on the local behavior of the image. $\bar{x}/2$ corresponds

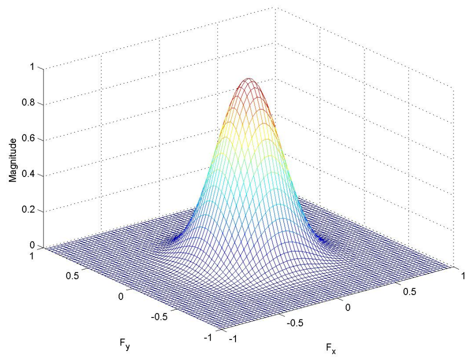

Figure 3: Transfer function of the Amacrine filter.

to half the mean value of the CFA pixel intensities over the whole CFA.

2.4 OPL processing

We assume that bipolar cells transmit the IPL signals to the OPL without any modification. We additionally suppose that amacrine cells operate similarly to horizontal cells. Thus, they act as a low pass filter on the bipolar signal and they modulate the adaptation parameter. We chose a 9x9 convolution filter having the transfer function given in Figure 3.

2.5 Demosaicing process

The final step of the processing is the demosaicing process. We apply a linear demosaicing method as described in [2]. Note that we do not apply any noise reduction algorithms. However, noise is amplified by the non-linear processing. In order to reduce the noise in the resulting image, we thus use a slightly different algorithm for demosaicing than described in [2]. We apply the following filters (Equation 3) for luminance and chrominance estimation in the CFA.

$$f = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} /16 \qquad f_{lum} = f * f \tag{3}$$
$$f_{chr} = \delta - (f * f) * (f * f) * f$$

where δ stands for the discrete Dirac function. We used bilinear interpolation to interpolate the chrominance. The low pass behaviors of the filters reduce the noise.

2.6 Results

We experimented with raw images from the digital camera Canon EOS30. We used the freeware tool dcraw.c[1] compiled under cygwin to extract a ppm image in 16-bits with the command line *dcraw -v -n -m -d -4 *.CRW*. The images are then processed with Matlab. The horizontal cell filtering is applied and that output is used to calculate the local parameter of the non-linear function. The next filtering, which is modeling the amacrine cells is applied, followed by demosaicing. A simple black and white point correction (histogram stretching) is done to render the image to display. Figure 4 shows an example of the method.

[1] http://www.cybercom.net/ dcoffin/dcraw/

Figure 4: Example of the method (a) The image is solely demosaiced (b) The first non-linearity and the demosaicing process (c) two non-linearities followed by demosaicing.

Figure 5: Comparison between several methods (a) method of [15] (b) method of [7] (c) our presented method.

Figure 6: Comparison between several methods (a) method in [15] (b) our method (c) method in [7] (d) our method.

Figure 5 and Figure 6 show a comparison of the proposed method with others algorithms. Additional results and comparisons are available on a our web site [1].

3. CONCLUSION

We defined a HDR rendering process that is directly applied on CFA images. The framework is inspired by the retinal processing that occurs in the human visual system. The method is fast and gives good results.

Our method loosely falls into the category of center-surround Retinex HDR algorithms [17]. As opposed to many of them, our method does not result in halo artifacts. We can avoid those without using an adaptive filter [15], which increases the computational speed tremendously.

Note that the presented method can be considered as a pre-processing method for tone mapping. Additional color rendering, such as white-balancing, color matricing, saturation correction and tone mapping needs to be applied for controlling the appearance for a specific color encoding or output device. For the figures in this article, we only applied a simple black and white point correction for rendering to display. Many of these corrections can be included inside the proposed processing pipeline by optimizing parameters, but this still needs to be demonstrated.

4. ACKNOWLEDGEMENT

This research was supported in part by the Swiss National Science Foundation (SNF) under grant No. 21-101681. We thank Frédo Durand for his comparison images.

REFERENCES

[1] Additional results are available at: http://ivrgwww.epfl.ch/misc/eusipco06/

[2] D. Alleysson, S. Süsstrunk, and J. Hérault, "Linear Demosaicing Inspired by the Human Visual System," *IEEE Transactions on Image Processing*, Vol. 14(4),pp. 439-449, 2005.

[3] D.A. Baylor, M.G.F. Fuortes, and P.M. O'Bryan, "Receptive fields of the cones in the retina of the turtle," *Journal of Physiology*, Vol. 214, pp. 265-294, 1971.

[4] D.M. Dacey, B.B. Lee, D.K. Stafford, J. Pokorny, and V.C. Smith, "Horizontal cells of the primate retina: cone specificity without spectral opponency," *Science*, Vol. 271, pp. 656-659, 1996.

[5] P.B. Detwiler and A.L. Hodgkin, "Electrical coupling between cones in the turtle retina," *Journal of Physiology*, Vol. 291, pp. 75-100, 1979.

[6] K. Devlin, "A review of tone reproduction techniques," *Department of Computer Science, University of Bristol, Tech. Rep. CSTR-02-005*, 2002.

[7] F. Durand and J. Dorsey, "Fast bilateral filtering for the display of high-dynamic-range images, *Proc. SIGGRAPH 02*, pp. 257-266, 2002.

[8] M.D. Fairchild and G.M. Johnson, "iCAM framework for image appearance differences and quality," *Journal of Electronic Imaging*, Vol. 13(1), pp. 126-138, 2004.

[9] R.D. Hamer and C.W. Tyler, "Phototransduction: Modeling the primate cone flash response," *Visual Neuroscience*, Vol. 12, pp. 1063-1082, 1995.

[10] C.R. Ingling and E. Martinez-Uriegas, "The spatiotemporal properties of the r-g x-cell channel", *Vision Research*, Vol. 25(1), pp. 33-38, 1985.

[11] H. Kolb, E. Fernandez, and R. Nelson, "WebVision: The organization of the Retina and Visual System", *http://webvision.med.utah.edu/*, Sept 2005.

[12] E. Land, "The Retinex," *American Scientist*, Vol. 52(2), pp. 247-264, 1964.

[13] E. Land and J. McCann, "Lightness and Retinex theory," *Journal of The Optical Society of America*, Vol. 61(1), pp. 1-11, 1971.

[14] P. Ledda, L. P. Santos, and A. Chalmers, "A local model of eye adaptation for high dynamic range images," *Proc. AFRIGRAPH2004*, ACM Press, pp. 151-160, 2004.

[15] L. Meylan and S. Süsstrunk, "High Dynamic Range Image Rendering Using a Retinex-Based Adaptive Filter," *to appear in IEEE Transactions on Image Processing*, 2006.

[16] O.S. Packer and D.M. Dacey, "Receptive field structure of H1 horizontal cells in macaque monkey retina," *Journal of Vision*, Vol. 2, pp. 272-292, 2002.

[17] Z. Rahman, D. D. Jobson, and G. A. Woodell, "Retinex processing for automatic image enhancement," *Journal of Electronic Imaging*, Vol. 13(1), pp. 100-110, 2004.

[18] E. Reinhard and K. Delvin, "Dynamic range reduction inspired by photoreceptor physiology," *IEEE Trans. On Visualization and Computer Graphics*, Vol. 11(1), pp. 13-24, 2005.

[19] A. Rizzi, C. Gatta, and D. Marini, "From Retinex to automatic color equalization issues in developing a new algorithm for unsupervised color equalization," *Journal of Electronic Imaging*, Vol. 13(1), pp. 15-28, 2004.

[20] A. Roorda and D.R. Williams, "The arrangement of the three cone classes in the living human eye," *Nature*, Vol. 397(11), pp. 520-522, 1999.

[21] R. Sobol, "Improving the Retinex algorithm for rendering wide dynamic range photographs," *Journal of Electronic Imaging*, Vol. 13(1), pp. 65-74, 2004.