Retinal Modeling in Digital Photography

Sabine Süsstrunk, Ecole Polytechnique Fédéral de Lausanne (EPFL), Lausanne, Switzerland;
joint work with David Alleysson, Psychology and NeuroCognition Laboratory, CNRS UMR 5105, Université Pierre-Mendes France (UPMF), Grenoble, France; Laurence Meylan (EPFL), and Daniel Tamburrino (EPFL)

Abstract

We present a digital camera workflow 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 non-linearities on the color signals 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. Our algorithm provides pleasing results on various images with different scene content, key, and dynamic range.

Many machine vision and image processing tasks can be solved more easily by considering how the human vision system processes information. This is especially the case for digital camera processing, where the goal usually is to output "pleasing" images. The term pleasing infers that an image resembles our perception (or memory) of the original scene, as opposed to the physical irradiance. How to measure perception in complex images, however, is a difficult task. While there is a tremendous amount of scientific literature, both in physiology and psychophysics, there are still too many gaps in our knowledge of visual information processing for engineers to build models and algorithms that consistently produce "high quality" images.

The part of the human visual system (HVS) best understood is the retina. We are well aware that there is a cone mosaic consisting of three different types of spectrally selective cones that capture light. The machine vision counterpart is the color filter array (CFA) placed in front of the sensor, which also captures different wavelengths at spatially distinct position given by the filter array pattern. Global adaptation to the scene's irradiances is implemented as a non-linear power-, sigmoid-, or log-function with black and white clipping. And the ganglion cells' receptive fields have their correspondence in luminance-chrominance color encodings.

In many imaging situations, especially if the scene dynamic range far exceeds the display dynamic range, modeling only global adaptation is not sufficient to render pleasing images. Thus, different algorithms have been proposed to model local adaptation, which effectively increases local contrast (see Figure 1). The best known are the computational Retinex alogrithms. The Retinex model of color vision, first formulated by Edwin Land, predicts that not absolute color responses, but response ratios calculated over scene colors remain constant. Subsequently, many algorithms were developed, their main difference being on how the ratios are calculated.

We took a different approach and designed a camera workflow based on a model of retinal processing. Contrary to other methods, we apply tone mapping directly on the mosaic image captured by the sensor, analogue to the human visual system that applies non-linearities on the color responses captured by the cone mosaic. In our retinal model, we assume that there are two adaptive non-linearities, one in the Outer Plexiform Layer (OPL) comprised of the cones, horizontal, and bipolar cells, and one in the Inner Plexiform Layer (IPL), which includes the bipolar, amacrine and ganglion cells. The non-linear function we apply is based on Naka-Rushton. The adaptation factor 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. Demosaicing is then applied to the tone mapped image.

Our algorithm is fast and provides pleasing results on various images with different scene content, key, and dynamic range. It can be used on raw images or as post-processing on already rendered images (see Figure).

Figure 1. *Top row: gamma encoded image (left), image processed with our workflow (right). Bottom row: original camera rendering (left), image post-processed with our workflow (right).*

The algorithms and results of our workflow are presented in references [1, 2, 3]. Code to implement our method can be found at [4].

References

[1] D. Alleysson, S. Süsstrunk, and J. Hérault, Linear demosaicing inspired by the human visual system, IEEE T IMAGE PROCESS, 14, 439 (2005).

[2] L. Meylan, D. Alleysson, and S. Süsstrunk, A Model of Retinal Local Adaptation for the Tone Mapping of Color Filter Array Images, J OPT SOC AM A, 24, 2807 (2007).

[3] D. Tamburrino, D. Alleysson, L. Meylan, and S. Süsstrunk, Digital Camera Workflow for High Dynamic Range Images Using a Model of Retinal Processing, IS&T/SPIE Proc. Digital Photography IV, 6817 (2008).

[4] http://ivrg.epfl.ch/supplementary_material/index.html

