

A Perceptually Inspired Variational Framework for Color Enhancement

Rodrigo Palma-Amestoy, Edoardo Provenzi,
Marcelo Bertalmío and Vicent Caselles *

December 1, 2025

Abstract

Basic phenomenology of human color vision has been widely taken as an inspiration to devise explicit color correction algorithms. The behavior of these models in terms of significative image features (such as contrast and dispersion) can be difficult to characterize. To cope with this, we propose to use a variational formulation of color contrast enhancement that is inspired by the basic phenomenology of color perception. In particular, we devise a set of basic requirements to be fulfilled by an energy to be considered as ‘perceptually inspired’, showing that there is an explicit class of functionals satisfying all of them. We single out three explicit functionals that we consider of basic interest, showing similarities and differences with existing models. The minima of such functionals is computed using a gradient descent approach. We also present a general methodology to reduce the computational cost of the algorithms under analysis from $\mathcal{O}(N^2)$ to $\mathcal{O}(N \log N)$, being N the number of input pixels.

1 Introduction

Human vision is a process of great complexity that involves many processes such as shape and pattern recognition, motion analysis and color perception. An essential literature about neurophysiological aspects of human vision can be found in the books [1–7].

*Rodrigo Palma-Amestoy is with Department of Electrical Engineering, Universidad de Chile, Avenida Tupper 2007, Casilla 412-3, Santiago, Chile. E-mail: ropalma@ing.uchile.cl. Tel. : +56 2 978 4207. Fax. +56 2 695 3881. Edoardo Provenzi is with Dipartimento di Tecnologia dell’Informazione, Università di Milano, Via Bramante 65, 26013, Crema (CR), Italy. E-mail: provenzi@dti.unimi.it. Tel. +39 0373 898089. Fax. +39 0373 898010. Marcelo Bertalmío and Vicent Caselles are with Departament de Tecnologia, Universitat Pompeu Fabra, Pg. de Circumvallació 8, 08003 Barcelona, Spain. E-mail: marcelo.bertalmio@upf.edu, Tel. +34 93 542 25 69, Fax. +34 93 542 25 17; E-mail: vicent.caselles@upf.edu, Tel. +34 93 542 24 21, Fax. +34 93 542 25 17.

In this paper we will focus on color perception. This process begins with light capture by the three different types of cone pigments inside the retina. They have different absorption characteristics, with peak absorptions in the red, green and blue regions of the optical spectrum. The existence of three types of cones provides a physiological basis for the trichromatic theory of color vision. When a light stimulus activates a cone, a photochemical transition occurs, producing a nerve impulse. Experimental evidence [7] show that the propagation of nerve impulses reach the brain, which analyzes and interprets the information carried by the impulses. The part of the brain devoted to color perception is the so-called V4 zone of the occipital cortex. Neither retina photochemistry nor nerve impulses propagation are well understood, hence a *deterministic* characterization of the visual process is still unavailable. For this reason, the majority of color perception models follow a *descriptive* approach, trying to simulate macroscopic characteristics of color vision, rather than reproduce neurophysiological activity.

The first modern contribution in such a field has been provided by the famous Retinex theory of Land and McCann [8]. They started devising a series of innovative experiments about color perception using unrealistic pictures called ‘Mondrians’ [9, 10]. The abstract character of the images minimized the action of high level faculties of the Human Visual System (HVS) such as memory learning, judgement and recognition, thus making possible the reproducibility of the experiments [11]. They then found a formal computation able to reproduce experimental data, providing the first Retinex model of color vision [8]. After that, a whole category of perceptually inspired computational algorithms has been developed with different aims: reproduction of color sensation, color stabilization, enhancement of color images, etc. [12]. A common characteristic of these works is that HVS features are taken as inspiration to devise the explicit equations of perceptually based algorithms.

In this paper, rather than dealing directly with explicit algorithms, we are interested in the analysis of the interplay between color perception and variational image processing. The reason is that variational techniques have proven to be a powerful instrument to both *understanding* and *solving* image processing problems. Variational principles are based on energy functionals, whose selection depends on the particular problem under examination. The image for which the functional reaches a minimum is the solution to the particular image processing task at hand. Since our aim is performing a perceptually inspired color correction, we are interested in the analysis of energy functionals complying with the basic features of color perception. So, we will provide a translation of basic macroscopic *phenomenological characteristics* of color vision into mathematical requirements to be fulfilled by an energy functional in order to be considered *perceptually inspired*. Then, the minimization of any such functional starting from a given input image will produce an output image which will be a color-corrected version of the original, and this correction process will have followed (some) basic behavior of the HVS.

We will show that it is possible to identify a class of functionals that comply with our set of basic perceptually based assumptions. Remarkably, this class happens to be the unique class that fulfills the logarithmic response to light stimuli expressed by Weber-Fechner’s law. The Euler-Lagrange equations corresponding to the minimization of these functionals can be used to implement computational (gradient descent) algorithms. The fact that these

equations come from a variational method permits to understand the algorithm’s behavior in terms of contrast enhancement and control of dispersion (departure from original gray levels and average value.) We will present a theoretical and qualitative analysis of three particular examples of perceptually inspired energy functionals that we consider of basic interest. This analysis will put in evidence similarities and differences with Retinex and a more recent perceptual color correction algorithm called ACE [13]. Furthermore, we will discuss a general procedure to reduce the computational cost of the algorithms derived by the three selected energy functionals from $\mathcal{O}(N^2)$ to $\mathcal{O}(N \log N)$, being N the number of pixels in the input image.

Let us finally describe the plan of the paper. In Section II we recall the basic HVS phenomenological properties that will lead us to formulate, in Section III, our basic set of assumptions that must be satisfied by an energy functional to be considered perceptually-inspired. Section IV contains three explicit examples of perceptual energies, the mathematical analysis of their variations is given in Section V. In Section VI we present a qualitative comparison between the algorithms corresponding to these energies and already existing perceptually-inspired algorithms. In Section VII we expose a general strategy to reduce the computational algorithm complexity. Section VIII contains the tests we have performed. We end with conclusions and open problems. All proofs have been put in the appendices.

2 Basic HVS phenomenological properties

The phenomenological properties of human color perception that we consider as basics for our axiomatic construction are three: color constancy, local contrast enhancement and visual adaptation. For the sake of clarity, we discuss them in different subsections.

2.1 Human color constancy

It is well known that the HVS has strong ‘normalization’ properties, i.e. humans can perceive colors of a scene almost independently of the spectral electromagnetic composition of a uniform illuminant, usually called color cast [8, 9]. This peculiar ability is known as *human color constancy* and it is not perfect [14, 15]. In fact it depends on several factors, such as cast strength, or amount of detail in the scene. In particular, if the variety of visual information is poor, the HVS can be deceived. In fact, with a suitable selection of the illuminant, two perfectly uniform surfaces with different reflectance can be perceived as having the same color. However, as proved by Land’s experiments [8], this cannot happen if the surface is not uniform, showing that human color perception strongly depends on the context, i.e. on the *relative* luminance coming from different places of an observed scene, rather than on the *absolute* luminance value of every single point. Nonetheless, being color constancy not complete, also absolute punctual luminance information plays a role in the entire color perception process and cannot be completely discarded.

The majority of color models try to remove the color cast due to illuminant looking for the invariant component of the light signal: the physical reflectance of objects (see [16]

for a recent overview of color constancy algorithms). However it is well known that the separation between illuminant and reflectance is an ill-posed problem, unless one imposes further constraints which are not verified by all images [17]. We will see later that we intend to remove color cast using another approach based on contrast enhancement. To understand why we relate color cast and contrast enhancement, we notice that an image with color cast is always characterized by one chromatic channel with remarkably different standard deviation with respect to the others (not to be confused with an image with dominant color, e.g. the close up of a leaf, in which is the average value of a channel that prevails). To have a quantitative example let us consider, for instance, the picture with strong blue cast shown in fig. 1 with the relative R,G,B histograms. The average values of R,G,B intensities are, respectively, $\mu_R = 8.67$, $\mu_G = 14.79$, $\mu_B = 60.18$, while the standard deviations are $\sigma_R = 17.59$, $\sigma_G = 27.03$, $\sigma_B = 130.56$.

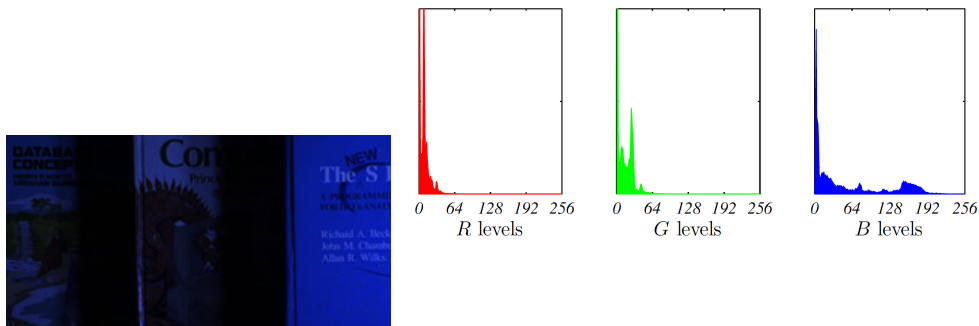


Figure 1: A picture with blue cast and its global R,G,B histograms.

An image without color cast does not show such a difference between standard deviations in the three chromatic channels. Being the standard deviation a measure of average contrast, it is clear that contrast enhancement can help decreasing the difference between σ_R , σ_G and σ_B , by spreading the intensity values of all separate chromatic channels, thus reducing color cast.

2.2 Local contrast enhancement

The HVS provides a local contrast enhancement, in the sense that the perception of details is spatially dependent. Well known phenomena exhibiting local contrast enhancement are Mach bands and simultaneous contrast [18]. Local HVS features related to color perception can be appreciated also in more complex scenes and go under the general name of ‘chromatic induction’ [19–22]. Experimental evidences show that the strength of chromatic induction between two different areas of a scene decreases monotonically with their Euclidean distance, even though a precise analytical description is not yet available.

2.3 Visual adaptation

The range of light intensity levels to which the HVS can adapt is impressive, on the order of 10^{10} , from the scotopic threshold to the glare limit [18]. However, the HVS system cannot operate over such a range simultaneously, rather it adapts to each scene to prevent saturation which would depress contrast sensitivity [23].

During this adaptation, the HVS *shifts* its sensitivity to light stimuli in order to present only modulations around the average intensity level of a scene [23]. This provides a phenomenological motivation for the so-called ‘gray-world’ (GW) assumption, which says that the average color in a scene is gray [24].

3 The basic set of assumptions

We now want to use the phenomenological characteristics just mentioned to write a set of basic assumptions to be satisfied by a perceptually inspired energy functional. Let us fix the notation that will be used throughout the paper. For notational simplicity, and also motivated by practical applications, we shall work with discrete images, even if many of the formulae we will consider could be written in an analogous context by replacing sums by integrals.

Given a discrete RGB image, we denote by $\mathfrak{I} = \{1, \dots, W\} \times \{1, \dots, H\} \subset \mathbb{Z}^2$ its spatial domain, $W, H \geq 1$ being integers; $x = (x_1, x_2)$ and $y = (y_1, y_2)$ denote the coordinates of two arbitrary pixels in $\mathfrak{I}$. We will always consider a normalized dynamic range in $[0, 1]$, so that a color image function is denoted by $\vec{I} : \mathfrak{I} \rightarrow [0, 1]^3$, $\vec{I}(x) = (I_R(x), I_G(x), I_B(x))$, where $I_k(x)$ is the intensity level of the pixel $x \in \mathfrak{I}$ in the chromatic channel $k \in \{R, G, B\}$. All computations will be performed on the scalar components of the image, thus treating independently each chromatic channel. As will be later explained with more detail, this choice permits us to properly treat images with color cast (see [2] chapter 8 ‘Color Vision’ for more details about independent chromatic channel computation and color cast).

For later convenience, we propose to use the following convention. Given any image $I : \mathfrak{I} \rightarrow [0, 1]$ we extend it as an even function with respect to the two variables in the domain $\{-W+1, \dots, W\} \times \{-H+1, \dots, H\}$ (which we still indicate with $\mathfrak{I}$ for simplicity), i.e. we replicate the image specularly in all directions, and then by periodicity to $\mathbb{Z} \times \mathbb{Z}$ with fundamental period $\mathfrak{I}$. For simplicity, we denote the extended image again by I . With this, we may consider the domain of I as the periodic sampling lattice, that is $\mathbb{T}_d := (\mathbb{Z} \times \mathbb{Z}) / (2W\mathbb{Z} \times 2H\mathbb{Z})$. This notation means that we identify any pair of points $x = (x_1, x_2)$ and $y = (y_1, y_2)$ in $\mathbb{Z} \times \mathbb{Z}$ if $x_1 - y_1 \in 2W\mathbb{Z}$ and $x_2 - y_2 \in 2H\mathbb{Z}$. We denote this equivalence relation by $\equiv$. The distance between any two points $x, y \in \mathbb{T}_d$, denoted by $\|x - y\|_{\mathbb{T}_d}$, is computed as $\min\{|\tilde{x} - \tilde{y}| : \tilde{x} \equiv x, \tilde{y} \equiv y\}$, where $|v| = \sqrt{v_1^2 + v_2^2}$, $v = (v_1, v_2)$. From now on, we shall assume that our images have these symmetry and are defined on the extended domain $\mathfrak{I}$.

3.1 The first assumption: General structure of the energy functional

In order to write the general structure of a perceptually inspired energy functional, we start noticing that human color perception is characterized by *both local and global features*: contrast enhancement has a local nature, i.e. *spatially variant*, while visual adaptation and attachment to original data (implied by the failure of color constancy) have a global nature, i.e. *spatially invariant*, in the sense that they do not depend on the intensity distribution in the neighborhood.

These basic considerations imply that a perceptually inspired energy functional should contain two terms: one spatially-dependent term whose minimization leads to a local contrast enhancement and one global term whose minimization leads to a control of the departure from both original punctual values and the middle gray, which, in our normalized dynamic range, is $1/2$.

Let us first describe the general form of the contrast enhancement terms that we will consider in this paper. For that we need a contrast measure $\bar{c}(a, b)$ between two gray levels $a, b > 0$ (to avoid some singular cases, we shall assume that intensity image values are always positive). We require the contrast function $\bar{c} : (0, +\infty) \times (0, +\infty) \rightarrow \mathbb{R}$ to be continuous, symmetric in (a, b) , i.e. $\bar{c}(a, b) = \bar{c}(b, a)$, increasing when $\min(a, b)$ decreases or $\max(a, b)$ increases. Basic examples of contrast measures are $\bar{c} = |a - b| \equiv \max(a, b) - \min(a, b)$ or $\bar{c}(a, b) = \frac{\max(a, b)}{\min(a, b)}$.

Since our purpose is to enhance contrast by *minimizing* an energy, we define an *inverse contrast function* $c(a, b)$, still continuous and symmetric in (a, b) , but decreasing when $\min(a, b)$ decreases or $\max(a, b)$ increases. Notice that, if $\bar{c}(a, b)$ is a contrast measure, then $c(a, b) = -\bar{c}(a, b)$ or $c(a, b) = 1/\bar{c}(a, b)$ is an inverse contrast measure, so that basic examples of inverse contrast are: $c(a, b) = \min(a, b) - \max(a, b)$ or $c(a, b) = \frac{\min(a, b)}{\max(a, b)}$.

Let us now introduce a weighting function to localize the contrast computation. Let $w : \mathfrak{I} \times \mathfrak{I} \rightarrow \mathbb{R}^+$ be a positive symmetric kernel, i.e. such that $w(x, y) = w(y, x) > 0$, for all $x, y \in \mathfrak{I}$, that measures the mutual influence between the pixels x, y . The symmetry requirement is motivated by the fact that the mutual chromatic induction is independent on the order of the two pixel considered. Usually, we assume that $w(x, y)$ is a function of the Euclidean distance $\|x - y\|_{\mathfrak{I}}$ between the two points. We shall assume that the kernel is normalized, i.e. that

$$\sum_{y \in \mathfrak{I}} w(x, y) = 1 \quad \forall x \in \mathfrak{I}. \quad (1)$$

Given an inverse contrast function $c(a, b)$ and a positive symmetric kernel $w(x, y)$, we define a contrast energy term by

$$C_w(I) = \sum_{x \in \mathfrak{I}} \sum_{y \in \mathfrak{I}} w(x, y) c(I(x), I(y)). \quad (2)$$

Thanks to the symmetry assumption, we may write $c(a, b) = \tilde{c}(\min(a, b), \max(a, b))$ for some function $\tilde{c}$ (indeed well defined by this identity). Notice that $\tilde{c}$ is non-decreasing

in the first argument and non-increasing in the second one. The symmetry hypothesis is not restrictive, in fact, if the inverse contrast measure $c(a, b)$ were not symmetric, we could write it as the sum $c(a, b) = c_s(a, b) + c_{as}(a, b)$ where $c_s(a, b)$ and $c_{as}(a, b)$ are symmetric and anti-symmetric respectively. Since the sum $\sum_{(x,y) \in \mathcal{I}^2} w(x, y) c_{as}(I(x), I(y)) = 0$, then the only remaining term is $\sum_{(x,y) \in \mathcal{I}^2} w(x, y) c_s(I(x), I(y))$, hence we may assume that $c(a, b)$ is symmetric in (a, b) .

Let us now consider the term that should control the dispersion. As suggested previously, we intend it as an attachment term to the initial given image I_0 and to the average illumination value, which we assume to be $1/2$. Thus, we define two dispersion functions: $d_1(I(x), I_0(x))$ to measure the separation between $I(x)$ and $I_0(x)$, and $d_2(I(x), \frac{1}{2})$ which measures the separation from the value $1/2$. Both d_1 and d_2 are continuous functions $d_{1,2} : \mathbb{R}^2 \rightarrow \mathbb{R}^+$ such that $d_{1,2}(a, a) = 0$ for any $a \in \mathbb{R}$, and $d_{1,2}(a, b) > 0$ if $a \neq b$. We write $d_{I_0, \frac{1}{2}}(I(x)) = d_1(I(x), I_0(x)) + d_2(I(x), \frac{1}{2})$, and the dispersion energy term as

$$D(I) = \sum_{x \in \mathcal{I}} d_{I_0, \frac{1}{2}}(I(x)). \quad (3)$$

We can now formulate our first assumption.

Assumption 1. The general structure of a perceptually inspired color correction energy functional is

$$E_w(I) = D(I) + C_w(I), \quad (4)$$

where $C_w(I)$ and $D(I)$ are the contrast and dispersion terms defined in (2) and (3), respectively. The minimization of D must provide a control of the dispersion around $1/2$ and around the original intensity values. The minimization of C_w must provide a local contrast enhancement.

We notice that, even if we have derived the above structure of a perceptual energy functional starting from phenomenological evidence, it agrees with the general structure proposed in [25] on the basis of neurophysiological evidences.

3.2 The second assumption: Properties of the contrast function

In order to find out which properties the contrast term should satisfy, let us observe that an overall change in intensity, measured by the generic quantity $\lambda > 0$, does not affect the visual sensation. This requires the contrast function c to be homogeneous, recalling that c is homogeneous of degree $n \in \mathbb{Z}$ if

$$c(\lambda a, \lambda b) = \lambda^n c(a, b) \quad \forall \lambda, a, b \in (0, +\infty), \quad (5)$$

where a and b are synthetic representations of $I(x)$ and $I(y)$. Of course, if $n = 0$, c automatically disregards the presence of λ , but we can say more: since λ can take any positive value, if we set $\lambda = 1/b$, we may write equation (5) as:

$$c(a, b) = b^n c\left(\frac{a}{b}, 1\right) \quad \forall a, b \in (0, +\infty), \quad (6)$$

so, when $n = 0$, $b^n = 1$ and c results as a function of the ratio a/b which intrinsically disregards overall changes in light intensity. If $n > 0$, then λ has a global influence and could be removed performing a suitable normalization (for instance, dividing by the n -th power of the highest intensity level). We can formalize these considerations in our second assumption.

Assumption 2. We assume that the inverse contrast function $c(a, b)$ is homogeneous.

Thanks to the arguments presented so far, we have that inverse contrast functions which are homogeneous of degree $n = 0$ are those that can be written as a monotone non-decreasing function of $\frac{\min(I(x), I(y))}{\max(I(x), I(y))}$. It would be interesting to know if, given an inverse contrast function c , there are increasing functions $g, h : (0, +\infty) \rightarrow (0, +\infty)$ such that $g(c(h(a), h(b))) = \frac{a}{b}$, for any $0 < a < b$. Probably some other assumptions on c are required to obtain such a characterization, which means that, modulo a calibration of intensity values, $\frac{\min(I(x), I(y))}{\max(I(x), I(y))}$ is the essential inverse contrast function. In the next subsection we provide further support to restrict ourselves to such class of contrast functions.

3.3 Contrast functions satisfying the Weber-Fechner's law

Weber-Fechner's law describes a common behavior of human senses. It has been first discovered by Weber and then formalized by Fechner. In general, it describes a non-linear response of a human sense to variations of external stimuli: the same variation is perceived in a weaker way as the strength of the external stimulus increases. The experiment for human vision is performed in controlled conditions [18]: on a uniform background of intensity I_0 , it is superimposed a brief pulse of light of intensity $I_1 > I_0$. Weber-Fechner's law states that the so-called *Weber-Fechner ratio* $\mathcal{R}_{WF} \equiv \frac{I_1 - I_0}{I_0}$, i.e. the ratio between the variation $\Delta I \equiv I_1 - I_0$ and the background I_0 , remains constant. Even though Weber-Fechner's law is not perfect [4], the intensity range over which it is in good agreement with experience (called 'Weber-Fechner's domain') is still comparable to the dynamic range of most electronic imaging systems [4].

Since $\mathcal{R}_{WF} = I_1/I_0 - 1$, Weber-Fechner's law is saying that the perceived contrast is a function of I_1/I_0 . This reason motivates us to say that $c(a, b)$ is a *generalized Weber-Fechner contrast function* if c is an *inverse contrast function* which can be written as a non-decreasing function of $\min(a, b)/\max(a, b)$. Hence, we can particularize assumption 2 as follows.

Assumption 2'. We assume that c is a generalized Weber-Fechner contrast function.

4 Explicit energy functionals complying with our set of assumptions

The purpose of this section is to show explicit examples of energy functionals complying with our set of assumptions. We first concentrate on the contrast terms, then we discuss the dispersion terms and, finally, we analyze the Euler-Lagrange equations of the complete energy functionals. Here we mainly concentrate on mathematical features, reserving qualitative interpretations and comparisons for the following section.

4.1 Three basic inverse contrast terms

As we have explained in Section 3 we shall concentrate on generalized Weber-Fechner contrast terms. The following three are of particular interest:

$$C_w^{\text{id}}(I) := \frac{1}{4} \sum_{x \in \mathfrak{I}} \sum_{y \in \mathfrak{I}} w(x, y) \frac{\min(I(x), I(y))}{\max(I(x), I(y))}, \quad (7)$$

$$C_w^{\log}(I) := \frac{1}{4} \sum_{x \in \mathfrak{I}} \sum_{y \in \mathfrak{I}} w(x, y) \log \left(\frac{\min(I(x), I(y))}{\max(I(x), I(y))} \right), \quad (8)$$

$$C_w^{-\mathcal{M}}(I) := -\frac{1}{4} \sum_{x \in \mathfrak{I}} \sum_{y \in \mathfrak{I}} w(x, y) \mathcal{M} \left(\frac{\min(I(x), I(y))}{\max(I(x), I(y))} \right), \quad (9)$$

where

$$\mathcal{M} \left(\frac{\min(I(x), I(y))}{\max(I(x), I(y))} \right) := \frac{1 - \frac{\min(I(x), I(y))}{\max(I(x), I(y))}}{1 + \frac{\min(I(x), I(y))}{\max(I(x), I(y))}} \equiv \frac{|I(x) - I(y)|}{I(x) + I(y)}, \quad (10)$$

is the well known Michelson's definition of contrast [26]. The upper symbol in the above definitions of C_w simply specifies the monotone function applied on the basic contrast variable $t := \frac{\min(I(x), I(y))}{\max(I(x), I(y))}$, i.e. identity $\text{id}(t) = t$, logarithm $\log t$, and minus the Michelson's contrast function $-\mathcal{M}(t) = -\frac{1-t}{1+t}$, respectively. Indeed, these three functions are strictly increasing (this is obvious for the first two, for the third observe that $-\mathcal{M}'(t) = 2/(1+t)^2 > 0$ for all t). Notice that the function $t = \min(I(x), I(y))/\max(I(x), I(y))$ is minimized when $\min(I(x), I(y))$ takes the smallest possible value and $\max(I(x), I(y))$ takes the largest possible one, which corresponds to a contrast stretching. Thus, minimizing an increasing function of the variable t , will produce a contrast enhancement. To refer to any one of them we use the notation $C_w^f(I)$, where $f = \text{id}, \log, -\mathcal{M}$.

Let us briefly comment on the reasons for considering those contrast terms among others in the Weber-Fechner class. The first one is the simplest, and it is useful to understand the basic contrast enhancement properties of our energy functionals. The second has been chosen because of the logarithmic property to transform ratios in differences, so that the basic contrast variable becomes $\min(\tilde{I}(x), \tilde{I}(y)) - \max(\tilde{I}(x), \tilde{I}(y)) = -|\tilde{I}(x) - \tilde{I}(y)|$, being $\tilde{I} \equiv \log I$. Finally, the third one has been chosen because it is given in terms of the integral of Michelson's contrast.

4.2 Entropic dispersion term

The main features of the dispersion term have to be its attachment to the initial given image I_0 and to the average illumination value, which we assume to be $1/2$. In principle, to measure the dispersion of I with respect to I_0 or $1/2$, any distance function can be used. The simplest example would be a quadratic distance

$$D_{\alpha,\beta}^q(I) := \frac{\alpha}{2} \sum_{x \in \mathcal{I}} \left(I(x) - \frac{1}{2} \right)^2 + \frac{\beta}{2} \sum_{x \in \mathcal{I}} (I(x) - I_0(x))^2, \quad \alpha, \beta > 0. \quad (11)$$

Given that contrast terms are expressed as homogeneous functions of degree 0, the variational derivatives are homogeneous functions of degree -1 (see Appendix VIII-C). Since our axioms do not give any precise indication about the analytical form of the dispersion term that should be chosen, we search for functions able to maintain coherence with this homogeneity. A good candidate for this is the *entropic dispersion* term, i.e.

$$D_{\alpha,\beta}^{\mathcal{E}}(I) := \alpha \sum_{x \in \mathcal{I}} \left(\frac{1}{2} \log \frac{1}{2I(x)} - \left(\frac{1}{2} - I(x) \right) \right) + \beta \sum_{x \in \mathcal{I}} \left(I_0(x) \log \frac{I_0(x)}{I(x)} - (I_0(x) - I(x)) \right), \quad (12)$$

where $\alpha, \beta > 0$, which is based on the relative entropy distance [27] between I and $1/2$ (the first term) and between I_0 and I (the second term). Notice that, if $a > 0$ and $f(s) = a \log \frac{a}{s} - (a - s)$, $s \in (0, 1]$, then $\frac{df}{ds}(s) = 1 - \frac{a}{s}$ and $\frac{d^2f}{ds^2}(s) = \frac{a}{s^2} > 0$, $\forall s$. So, $f(s)$ has a global minimum in $s = a$. In particular, this holds when $a = I_0(x)$ or $a = 1/2$. Given the statistical interpretation of entropy, we can say that *minimizing $D_{\alpha,\beta}^{\mathcal{E}}(I)$ amounts to minimizing the disorder of intensity levels around $1/2$ and around the original data $I_0(x)$* . Thus, $D_{\alpha,\beta}^{\mathcal{E}}(I)$ accomplishes the required tasks of a dispersion term.

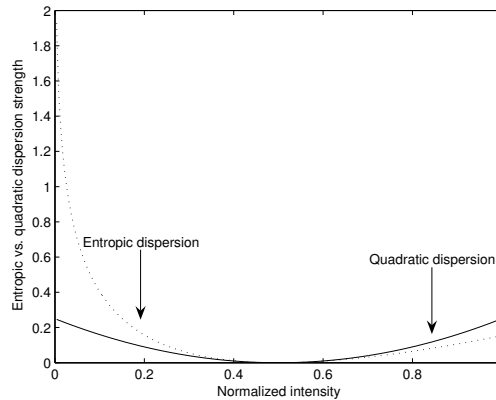


Figure 2: *Dotted line.* Graphic of entropic dispersion function around $1/2$. *Solid line.* Graphic of quadratic dispersion function around $1/2$.

Notice that the integrand functions of the entropic dispersion term are not symmetric with respect to their minima, as instead happens for the ones of the quadratic dispersion

term $D_{\alpha,\beta}^q(I)$. The difference can be appreciated in the graphics of Fig.2 which refer, for simplicity, to the dispersion around $1/2$.

In the following section we show the explicit minimization of these energy functionals.

5 Minimization of the energy functionals $E_{w,\alpha,\beta}^f(I) = D_{\alpha,\beta}^\mathcal{E}(I) + C_w^f(I)$

Recall that the minimization of $E_{w,\alpha,\beta}^f(I)$, $f = \text{id}, \log, -\mathcal{M}$ corresponds to a trade-off between two opponent mechanisms: on one hand we have entropic control of dispersion around $1/2$ and around original data, on the other hand we have local contrast enhancement.

The existence of a minimum in the discrete framework can be guaranteed for a quite general class of energy functionals. Let us define, for $\rho > 0$, $\mathcal{F}_\rho := \{I : \mathfrak{I} \rightarrow [0, 1], I(x) \geq \rho \forall x \in \mathfrak{I}\}$. We minimize the energy E in the class $\mathcal{F}_\rho$ to avoid singularities when we use $f = \log$ or in the entropic dispersion term.

Proposition 5.1 *Let $c : (0, 1] \times (0, 1] \rightarrow \mathbb{R}$ be a continuous function. Let $E(I) := D_{\alpha,\beta}^\mathcal{E}(I) + C_w(I)$ where $C_w(I)$ is given in (2). There is a minimum of $E(I)$ in the class of functions $\mathcal{F}_\rho$. An analogous result holds if we use the quadratic dispersion term $D_{\alpha,\beta}^q(I)$.*

The proof is given in Appendix VIII-A.

Notice that if the energy E is differentiable the minimum I^* satisfies $\delta E(I^*) = 0$. Before computing the Euler-Lagrange equations, we notice that the contrast terms $C_w^f(I)$, $f = \text{id}, \log, -\mathcal{M}$ are not convex and the basic function $t := \frac{\min(a,b)}{\max(a,b)}$ is not differentiable. In fact, we may write $\min(a,b) = \frac{1}{2}(a+b-|a-b|)$, $\max(a,b) = \frac{1}{2}(a+b+|a-b|)$, for any $a, b \in \mathbb{R}$. The non-differentiability comes from the absolute value $A(z) = |z|$, $z \in \mathbb{R}$. Since our algorithm will use a gradient descent approach, we must regularize the basic variable t . We notice that $A'(z) = 1$ if $z > 0$, $A'(z) = -1$ if $z < 0$ and A is not differentiable at $z = 0$. But all the values $s \in [-1, 1]$ are *subtangents* of $A(z)$ at $z = 0$, that is, $A(z) - A(0) \geq s(z - 0)$ for any $z \in \mathbb{R}$. Thus we may write $A'(z) = \text{sign}(z)$, where

$$\text{sign}(z) = \begin{cases} 1 & \text{if } z > 0 \\ [-1, 1] & \text{if } z = 0 \\ -1 & \text{if } z < 0 \end{cases} . \quad (13)$$

We define $\text{sign}_0(z)$ as in (13), but with the particular choice 0 when $z = 0$.

Def. 5.1 *Given $\epsilon > 0$, we say that $A_\epsilon(z)$ is a ‘nice regularization’ of $A(z)$, if $A_\epsilon(z) \geq 0$ is convex, differentiable with continuous derivative, $A_\epsilon(0) = 0$, $A_\epsilon(-z) = A_\epsilon(z)$, and*

- (i) $A_\epsilon(z) \leq |z|$ for any $z \in \mathbb{R}$ and $A_\epsilon(z) = |z| + Q_{1,\epsilon}(z)$ where $Q_{1,\epsilon}(z) \rightarrow 0$ as $\epsilon \rightarrow 0$, uniformly in $z \in [-1, 1]$;
- (ii) Let us denote $s_\epsilon(z) = A'_\epsilon(z)$. Then $|s_\epsilon(z)| \leq 1$ for any $z \in [-1, 1]$, $s_\epsilon(z) \rightarrow \text{sign}_0(z)$ as $\epsilon \rightarrow 0$ for any $z \in \mathbb{R}$, and $Q_{2,\epsilon}(z) := A_\epsilon(z) - zs_\epsilon(z) \rightarrow 0$ as $\epsilon \rightarrow 0$, uniformly in $z \in [-1, 1]$.

We present two examples of nice regularization of $A(z)$. *Example a):* $A_\epsilon(z) = \sqrt{\epsilon^2 + |z|^2} - \epsilon$, in this case $s_\epsilon(z) = \frac{z}{\sqrt{\epsilon^2 + |z|^2}}$, $Q_{1,\epsilon}(z) = O(\epsilon)$ and $Q_{2,\epsilon}(z) := A_\epsilon(z) - zs_\epsilon(z) = O(\epsilon)$.

Example b): $A_\epsilon(z) = z \frac{\arctan(z/\epsilon)}{\arctan(1/\epsilon)} - \frac{\epsilon}{2 \arctan(1/\epsilon)} \log(1 + \frac{z^2}{\epsilon^2})$, in this case $s_\epsilon(z) = \frac{\arctan(z/\epsilon)}{\arctan(1/\epsilon)}$, $Q_{1,\epsilon}(z) = O(\epsilon \log(1/\epsilon))$ and $Q_{2,\epsilon}(z) = O(\epsilon \log(1/\epsilon))$, uniformly in $z \in [-1, 1]$. We have denoted by $O(F(\epsilon))$ any expression satisfying $|O(F(\epsilon))| \leq CF(\epsilon)$ for some constant $C > 0$ and $\epsilon > 0$ small enough. Observe that, in both cases $s_\epsilon(z) \rightarrow \text{sign}_0(z)$ as $\epsilon \rightarrow 0$ for any $z \in \mathbb{R}$. We give the details of the above statements in Appendix VIII-B.

Now let us assume that $A_\epsilon(z)$ is a nice regularization of $A(z)$. We set

$$\min_\epsilon(a, b) = \frac{1}{2}(a + b - A_\epsilon(a - b)), \quad \max_\epsilon(a, b) = \frac{1}{2}(a + b + A_\epsilon(a - b)). \quad (14)$$

We define the regularized version of the functionals as:

$$C_{w,\epsilon}^{\text{id}}(I) := \frac{1}{4} \sum_{x \in \mathfrak{J}} \sum_{y \in \mathfrak{J}} w(x, y) \frac{\min_\epsilon(I(x), I(y))}{\max_\epsilon(I(x), I(y))}; \quad (15)$$

$$C_{w,\epsilon}^{\log}(I) := \frac{1}{4} \sum_{x \in \mathfrak{J}} \sum_{y \in \mathfrak{J}} w(x, y) \log \left(\frac{\min_\epsilon(I(x), I(y))}{\max_\epsilon(I(x), I(y))} \right); \quad (16)$$

$$C_{w,\epsilon}^{-\mathcal{M}}(I) := -\frac{1}{4} \sum_{x \in \mathfrak{J}} \sum_{y \in \mathfrak{J}} w(x, y) \frac{A_\epsilon(I(x) - I(y))}{I(x) + I(y)}. \quad (17)$$

Proposition 5.2 Assume that $A_\epsilon(z)$ is a nice regularization of $A(z)$.

(i) The first variation of $C_{w,\epsilon}^{\text{id}}(I)$ is:

$$\begin{aligned} \delta C_{w,\epsilon}^{\text{id}}(I) &= -\frac{1}{2} \sum_{y \in \mathfrak{J}} w(x, y) \frac{I(y)}{\max_\epsilon(I(x), I(y))^2} s_\epsilon(I(x) - I(y)) + S_\epsilon \\ &= -\frac{1}{2} \sum_{y \in \mathfrak{J}} w(x, y) \frac{I(y)}{\max(I(x), I(y))^2} s_\epsilon(I(x) - I(y)) + S'_\epsilon, \end{aligned} \quad (18)$$

where $S_\epsilon, S'_\epsilon = O(Q_{1,\epsilon}(I(x) - I(y)) + Q_{2,\epsilon}(I(x) - I(y)))$; notice the difference between $\max$ and $\max_\epsilon$ in the previous formula;

(ii) The first variation of $C_{w,\epsilon}^{\log}(I)$ is:

$$\delta C_{w,\epsilon}^{\log}(I) = -\frac{1}{2} \sum_{y \in \mathfrak{J}} w(x, y) \frac{1}{I(x)} s_\epsilon(I(x) - I(y)) + S_\epsilon, \quad (19)$$

where $S_\epsilon = O(Q_{1,\epsilon}(I(x) - I(y)) + Q_{2,\epsilon}(I(x) - I(y)))$;
(iii) The first variation of $C_{w,\epsilon}^{-\mathcal{M}}(I)$ is:

$$\delta C_{w,\epsilon}^{-\mathcal{M}}(I) = - \sum_{y \in \mathcal{I}} w(x, y) \frac{I(y)}{(I(x) + I(y))^2} s_\epsilon(I(x) - I(y)) + S_\epsilon, \quad (20)$$

where $S_\epsilon = O(Q_{2,\epsilon}(I(x) - I(y)))$.

In all cases $Q_{1,\epsilon}(z), Q_{2,\epsilon}(z) \rightarrow 0$ as $\epsilon \rightarrow 0$, uniformly in $z \in [-1, 1]$.

Thus we know that $S_\epsilon = O(\epsilon)$ if $A_\epsilon(z)$ is given by Example a), and $S_\epsilon = O(\epsilon \log 1/\epsilon)$ if $A_\epsilon(z)$ is given by Example b). These are the cases of interest for us in the experiments.

The proof of this proposition will be given in Appendix VIII-C.

Notice that, letting $\epsilon \rightarrow 0$ we have that $\delta C_{w,\epsilon}^f(I) \rightarrow \delta C_{w,0}^f(I)$, where

$$\begin{aligned} \delta C_{w,0}^{\text{id}}(I) &= -\frac{1}{2} \sum_{y \in \mathcal{I}} w(x, y) \frac{I(y)}{\max(I(x), I(y))^2} \text{sign}_0(I(x) - I(y)) \\ &= -\frac{1}{2} \left(\sum_{\{y \in \mathcal{I} : I(x) > I(y)\}} w(x, y) \frac{I(y)}{I(x)^2} - \sum_{\{y \in \mathcal{I} : I(x) < I(y)\}} w(x, y) \frac{1}{I(y)} \right); \end{aligned}$$

$$\begin{aligned} \delta C_{w,0}^{\log}(I) &= -\frac{1}{2} \sum_{y \in \mathcal{I}} w(x, y) \frac{1}{I(x)} \text{sign}_0(I(x) - I(y)) \\ &= -\frac{1}{2} \left(\sum_{\{y \in \mathcal{I} : I(x) > I(y)\}} w(x, y) \frac{1}{I(x)} - \sum_{\{y \in \mathcal{I} : I(x) < I(y)\}} w(x, y) \frac{1}{I(x)} \right); \end{aligned}$$

$$\begin{aligned} \delta C_{w,0}^{-\mathcal{M}}(I) &= - \sum_{y \in \mathcal{I}} w(x, y) \frac{I(y)}{(I(x) + I(y))^2} \text{sign}_0(I(x) - I(y)) \\ &= - \sum_{\{y \in \mathcal{I} : I(x) > I(y)\}} w(x, y) \frac{I(y)}{(I(x) + I(y))^2} + \sum_{\{y \in \mathcal{I} : I(x) < I(y)\}} w(x, y) \frac{I(y)}{(I(x) + I(y))^2}. \end{aligned}$$

Now, by direct computation, we have that the derivative of the entropic dispersion term is:

$$\delta D_{\alpha,\beta}^{\mathcal{E}}(I) = \alpha \left(1 - \frac{1}{2I(x)} \right) + \beta \left(1 - \frac{I_0(x)}{I(x)} \right). \quad (21)$$

We can see that this expression has a degree of homogeneity -1 with respect to $I(x)$, the same as the variation of the three contrast terms $C_w^f(I)$.

Assume that $\alpha, \beta > 0$ are fixed. If $E_{w,\epsilon}^f(I) = D_{\alpha,\beta}^\mathcal{E}(I) + C_{w,\epsilon}^f(I)$, $f = \text{id}, \log, -\mathcal{M}$, then by linearity of the variational derivative, we have $\delta E_{w,\epsilon}^f(I) = \delta D_{\alpha,\beta}^\mathcal{E}(I) + \delta C_{w,\epsilon}^f(I)$. The minimum of $E_{w,\epsilon}^f(I)$ satisfies $\delta E_{w,\epsilon}^f(I) = 0$. To search for the minimum we use a semi-implicit discrete gradient descent strategy with respect to $\log I$. The continuous gradient descent equation is

$$\partial_t \log I = -\delta E_{w,\epsilon}^f(I), \quad (22)$$

being t the evolution parameter. Since $\partial_t \log I = \frac{1}{I} \partial_t I$, we have

$$\partial_t I = -I \delta E_{w,\epsilon}^f(I). \quad (23)$$

Using the gradient descent in $\log I$ leads to (23), which is related to a gradient descent approach which uses *the relative entropy as a metric*, instead of the usual quadratic distance (see [27]).

Let us now discretize our scheme: choosing a finite evolution step $\Delta t > 0$ and setting $I^k(x) = I_{k\Delta t}(x)$, $k = 0, 1, 2, \dots$, being $I^0(x)$ the original image, thanks to (21), we can write the semi-implicit discretization of (23) as

$$\frac{I^{k+1}(x) - I^k(x)}{\Delta t} = \alpha \left(\frac{1}{2} - I^{k+1}(x) \right) + \beta (I_0(x) - I^{k+1}(x)) - I^k(x) \delta C_{w,\epsilon}^f(I^k). \quad (24)$$

Now, considering the explicit expressions of $\delta C_{w,\epsilon}^f(I^k)$, neglecting their second terms containing S_ϵ (see Proposition 5.2 (i),(ii),(iii)) and performing some easy algebraic manipulations, we find the equation

$$I^{k+1}(x) = \frac{I^k(x) + \Delta t \left(\frac{\alpha}{2} + \beta I_0(x) + \frac{1}{2} R_{\epsilon, I^k}^f(x) \right)}{1 + \Delta t(\alpha + \beta)}, \quad (25)$$

where the function $R_{\epsilon, I^k}^f(x)$ assumes three different forms for $f = \text{id}, \log, -\mathcal{M}$, precisely

$$R_{\epsilon, I^k}^{\text{id}}(x) := -2I^k \delta C_{w,\epsilon}^{\text{id}}(I^k) = \sum_{y \in \mathcal{I}} w(x, y) \frac{I^k(x) I^k(y)}{\max_\epsilon(I^k(x), I^k(y))^2} s_\epsilon(I^k(x) - I^k(y)). \quad (26)$$

$$R_{\epsilon, I^k}^{\log}(x) := -2I^k \delta C_{w,\epsilon}^{\log}(I^k) = \sum_{y \in \mathcal{I}} w(x, y) s_\epsilon(I^k(x) - I^k(y)). \quad (27)$$

$$R_{\epsilon, I^k}^{\mathcal{M}}(x) := -2I^k \delta C_{w,\epsilon}^{-\mathcal{M}}(I^k) = \sum_{y \in \mathcal{I}} w(x, y) \frac{2I^k(x) I^k(y)}{(I^k(x) + I^k(y))^2} s_\epsilon(I^k(x) - I^k(y)). \quad (28)$$

At the limit $\epsilon \rightarrow 0$ we have:

$$R_{0, I^k}^{\text{id}}(x) = \sum_{\{y \in \mathcal{I} : I^k(x) > I^k(y)\}} w(x, y) \frac{I^k(y)}{I^k(x)} - \sum_{\{y \in \mathcal{I} : I^k(x) < I^k(y)\}} w(x, y) \frac{I^k(x)}{I^k(y)}; \quad (29)$$

$$R_{0,I^k}^{\log}(x) = \sum_{y \in \mathcal{I}} w(x, y) \text{sign}_0(I^k(x) - I^k(y)); \quad (30)$$

$$R_{0,I^k}^{-\mathcal{M}}(x) = \sum_{\mathcal{I}} w(x, y) \frac{2I^k(x)I^k(y)}{(I^k(x) + I^k(y))^2} \text{sign}_0(I^k(x) - I^k(y)). \quad (31)$$

In Appendix VIII-D, we will comment on the *stability* of these iterative methods.

6 Features of the three computational algorithms and comparison with existing models

Equation (25) can be used to implement three iterative computational algorithms for color image enhancement, as f varies. Notice that the R functions share an identical structure, i.e. $R_{\epsilon, I^k}^f(x) = \sum_{y \in \mathcal{I}} w(x, y) r_{\epsilon}^f(I^k(x), I^k(y))$, for a suitable function r_{ϵ}^f . In Fig. 3 we show the corresponding surfaces.

We can see that the surfaces representing r_{ϵ}^{id} and $r_{\epsilon}^{-\mathcal{M}}$ are very similar, the only difference being that the second is quite smoother than the first. Hence, we expect the corresponding algorithms to have similar enhancing properties.

Instead, the surface representing $r_{\epsilon}^{\log}$ shows an opposite behavior with respect to the previous two surfaces at the extreme points. Moreover, since the logarithmic function grows very rapidly for small values of its argument and slowly for higher ones, we expect the corresponding algorithm to perform a strong contrast enhancement in dark image areas and a weak one in bright zones. In section 8 it will be shown that all these observations are verified by empirical tests.

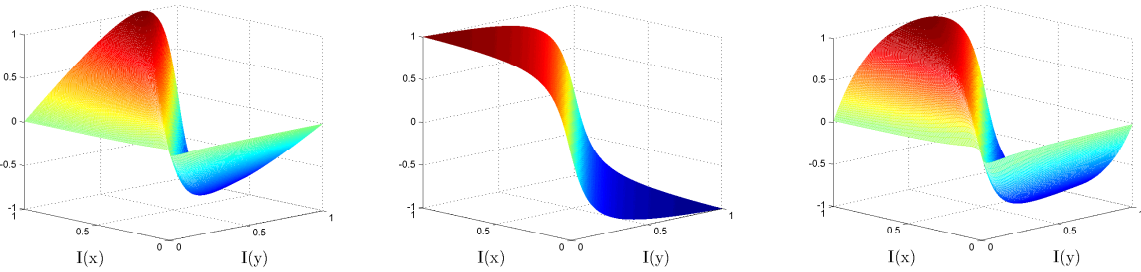


Figure 3: *Left:* Surface of r_{ϵ}^{id} . *Center:* Surface of $r_{\epsilon}^{\log}$. *Right:* Surface of $r_{\epsilon}^{-\mathcal{M}}$. Corresponding to the function $s_{\epsilon}(z) = \arctan(z/\epsilon)/\arctan(1/\epsilon)$ with $\epsilon = 1/20$.

We now want to show that there is a precise correspondence between the algorithm corresponding to $E_{w,\epsilon}^{\log}$ and ACE [13], a recent perceptually inspired color correction algorithm, while the algorithms corresponding to the energies $E_{w,\epsilon}^{\text{id}}$, $E_{w,\epsilon}^{-\mathcal{M}}$ are *qualitatively* related with the Retinex model, but they also exhibit new peculiarities that distinguish them from Retinex.

6.1 Correspondence between $E_{w,\epsilon}^{\log}$ and the ACE model

The original ACE model has been developed as a perceptually inspired color correction algorithm based on the already mentioned ‘gray-world principle’, i.e. that the average color of a scene is the middle gray [13]. The similarities between the ACE equations and those relative to variational histogram equalization [28] lead to a variational formulation of ACE and a deeper understanding of the model [29]. The functional energy of ACE is [29]:

$$E_w^{\text{ACE}}(I) := \sum_{x \in \mathfrak{I}} \left[\frac{\alpha}{2} \left(I(x) - \frac{1}{2} \right)^2 + \frac{\beta}{2} (I(x) - I_0(x))^2 \right] - \frac{1}{4} \sum_{x \in \mathfrak{I}} \sum_{y \in \mathfrak{I}} w(x, y) S(I(x) - I(y)), \quad (32)$$

being S the primitive function of a sigmoid $s: [-1, 1] \rightarrow [-1, 1]$ and $\alpha, \beta > 0$.

Now, applying a semi-implicit gradient descent technique to minimize E_w^{ACE} with respect to I , one arrives to the iterative scheme:

$$I^{k+1}(x) = \frac{I^k(x) + \Delta t \left(\frac{\alpha}{2} + \beta I_0(x) + \frac{1}{2} R_{I^k}^{\text{ACE}}(x) \right)}{1 + \Delta t(\alpha + \beta)}, \quad (33)$$

where

$$R_{I^k}^{\text{ACE}}(x) := \sum_{y \in \mathfrak{I}} w(x, y) s(I^k(x) - I^k(y)). \quad (34)$$

Since the sigmoid can be considered as an approximation of the sign function, the algorithms defined by (27) and (34) are essentially identical. We recall that the sigmoid in ACE has been used instead of the signum simply to avoid abrupt color corrections.

6.2 Qualitative similarities and differences between $E_{w,\epsilon}^{\text{id}}$, $E_{w,\epsilon}^{-\mathcal{M}}$ and the Retinex model

The models represented by the energies $E_{w,\epsilon}^{\text{id}}$, $E_{w,\epsilon}^{-\mathcal{M}}$ are, to the authors’ knowledge, novel in the literature. As we will show in the experiments in Section 8, the output images of their corresponding computational algorithms are very similar. Analytically (modulo replacing the function sign_ϵ by sign_0 for the sake of interpretation), this is motivated by the fact that both $R_{\epsilon, I^k}^{\text{id}}(x)$ and $R_{\epsilon, I^k}^{-\mathcal{M}}(x)$ are represented by integrals of intensity ratios bounded by 1 taken with positive sign when $I(x) > I(y)$ and negative sign in the opposite situation. The presence of ratios bounded by 1 is one of the features of the original Retinex model [8]. In fact, as it has been proven in [30], the peculiar ‘ratio-reset’ Retinex computation can be equivalently represented as the ratio between the intensity of any fixed pixel and the highest intensity sampled by a geometric structure used to extract spatial chromatic information around that pixel. The geometric structures used in the various Retinex implementations are many: paths, masks, sprays, etc. See, e.g., [31] for a more detailed overview.

The fundamental difference between Retinex and equations (26) and (28) is that, in the Retinex model, ratios are performed only over the highest intensity value and they are always taken with positive sign. As a consequence, as proved in [30], the ratio-reset

Retinex operation is unbalanced, i.e. it always increases the intensity of any pixel (a post-LUT processing can overcome this problem, but this operation is not intrinsically contained in the ratio-reset Retinex equations). Instead, equations (26) and (28) can both increase or reduce pixel intensity. For this reason, the energies $E_{w,\epsilon}^{\text{id}}$, $E_{w,\epsilon}^{-\mathcal{M}}$, could be *qualitatively* thought as describing a ‘symmetrized Retinex model’. Later we will compare the action of this model with the one of Retinex on an overexposed image, exhibiting the advantages provided by the symmetrization just commented. Finally, we would like to mention that a heuristic attempt to unify the qualities of Retinex and those of ACE in a single algorithm called RACE has been made in [32], equations (26) and (28) can be thought as a rigorous variational generalization of the RACE algorithm.

7 A general strategy for computational complexity reduction

The computational complexity of the algorithms derived from the minimization of the energy functionals that we have examined is $\mathcal{O}(N^2)$, N being the number of pixels of the input image. In fact, the computational complexity of $R_{\epsilon,I^k}^{\text{id}}(x)$, $R_{\epsilon,I^k}^{\log}(x)$ and $R_{\epsilon,I^k}^{-\mathcal{M}}(x)$ is of order $\mathcal{O}(N)$ for every x , hence the *global* computational complexity is of order $\mathcal{O}(N^2)$. This implies that with a standard PC (P4, 3GHz) it can take up hours to process a high resolution image.

In [31] a local sampling technique has been devised to reduce time computation. Here we provide a different technique, based on mathematical properties of the r -functions rather than on sampling. This permits to maintain all the chromatic information, avoiding sampling noise.

From now on, we assume that w is a function of $\|x - y\|_{\mathcal{J}}$ and we show how to speed up the computation of $R_{\epsilon,I^k}^f(x)$, reducing its computational complexity to $\mathcal{O}(N \log N)$ by means of a suitable approximation, which generalizes an analogous technique used in [29] to speed up the ACE algorithm. Let us denote with $R(x) = \sum_{y \in \mathcal{J}} w(x, y) r(I(x), I(y))$ any of the three R -functions specified in equations (26), (27), (28), omitting the indexes f , ϵ and I^k both to simplify the notation and because their specification is not important for what follows.

The basic observation underlying our proposal is that, if we manage to separate the dependence of the function r on $I(x)$ and $I(y)$, then we can express $R(x)$ as a sum of convolutions. In fact, let us suppose that $R(x)$ admits a polynomial approximation $R^{(n)}(x)$, being $n \in \mathbb{N}$ the approximation degree, such as

$$R^{(n)}(x) = \sum_{j=0}^n f_j(I(x)) \sum_{y \in \mathcal{J}} w(x, y) g_j(I(y)), \quad (35)$$

where f_j are functions of $I(x)$ and g_j are functions of $I(y)$. Since w is a function of $x - y$, it is clear that the integrals appearing in (35) actually reduce to convolutions, which can be computed using the Fast Fourier Transform (FFT), whose computational complexity is of

order $\mathcal{O}(N \log N)$. Hence, if $n \ll N$, the computational complexity of the approximated algorithm reduces to $\mathcal{O}(N \log N)$.

To achieve the separation of variables in (35), we fix an approximation degree n and we use a numerical algorithm to find the polynomial $p(I(x), I(y)) = \sum_{j=0}^n \sum_{l=0}^j p_{j-l,l} I(x)^{j-l} I(y)^l$, $p_{j-l,l} \in \mathbb{R}$, that minimizes the L^2 distance to $r(I(x), I(y))$. Indeed, since r is computed only once (before the first gradient descent iteration,) we can choose the degree n to control the maximum error. The polynomial can be easily rearranged and expressed as $p(I(x), I(y)) = \sum_{j=0}^n f_j(I(x)) I(y)^j$ for suitable functions $f_j(I(x))$. This expression can be directly used in equation (35) setting $g_j(I(y)) \equiv I^j(y)$ for every $j = 1, \dots, n$. Increasing n , one can reduce the approximation error; our experiments show that for the functions r_ϵ^{id} , r_ϵ^{log} and $r_\epsilon^{-\mathcal{M}}$, the value $n = 9$ gives a satisfactory approximation, so we have used it in all of our tests.

For the sake of completeness, we notice that, if the surface representing the function $r(I(x), I(y))$ has discontinuities or strong gradients, a polynomial expansion may not be useful for its approximation. In such a case, a Fourier expansion could be more suitable than the polynomial one, but the corresponding expression of $g_j(I(y))$ would be analytically more complicated, being, in this case, expressed in terms of trigonometric functions instead of powers.

8 Tests

So far, we have performed a theoretical investigation about the proposed axiomatic framework for perceptually inspired variational color enhancement. In this section, we show and discuss some tests about the algorithms derived from the three functional energies examined. We will work on images acquired with generic digital cameras. Before going into details, it is worthwhile to recall that the algorithms represented by equations (25), (26), (27) and (28) have four explicit parameters: α , β , ϵ and Δt , moreover, the function w determines the local contrast enhancement behavior. We have chosen the function $w(x, y) = \frac{A}{\|x-y\|_2}$, where A is the normalization parameter, so that (1) is satisfied. Recall that α controls the dispersion around the middle gray, β sets the strength of attachment to original data, and Δt is the gradient descent step. The parameter ϵ controls the regularization of the signum function. In practice, we have chosen $s_\epsilon(z) = \arctan(z/\epsilon)/\arctan(1/\epsilon)$ because the signum function is too singular to be used without leading to unwanted noise amplification.

Our tests indicate that an overall satisfactory set of parameters is $\alpha = \frac{255}{253}$, $\beta = 1$, $\Delta t = 0.2$, and $\epsilon = \frac{1}{20}$. We iterate until the average Mean Square Error (MSE) per pixel (between the images in the current and previous iteration) falls below a threshold of 10^{-4} (typically the steady state is reached within 10-20 iterations.) Varying these parameters with respect to image characteristics could possibly give better results. We consider the image-dependency of parameters an interesting issue to be investigated, but its thorough and significative analysis falls outside the scope of this paper.

8.1 Some experimental results

To show the abilities of the computational algorithms that we are discussing, we consider the three different images shown in fig. 4 (courtesy of P. Greenspun).

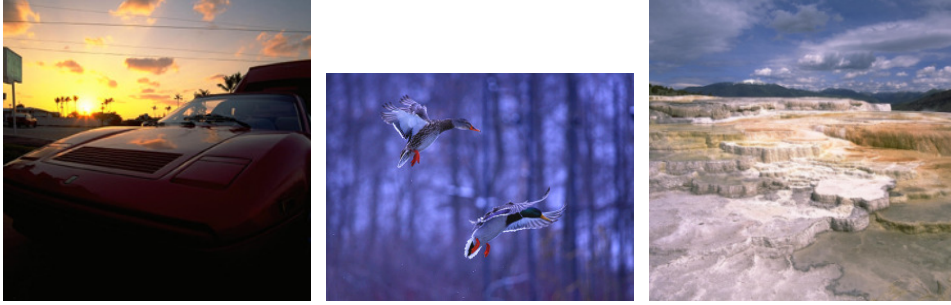


Figure 4: Three different original images of our test set showing different features to be enhanced.

The image on the left is underexposed, the image in the center has a strong blue cast, and the one on the right is quite overexposed. The results of the algorithms corresponding to equations (26), (27) and (28) are presented in figs. 5, 6 and 7, respectively.

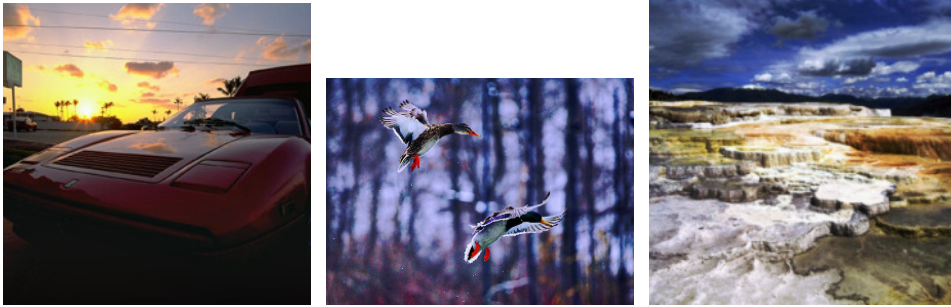


Figure 5: Outputs of the algorithm corresponding to the energy $E_{w,\epsilon}^{\text{id}}$.

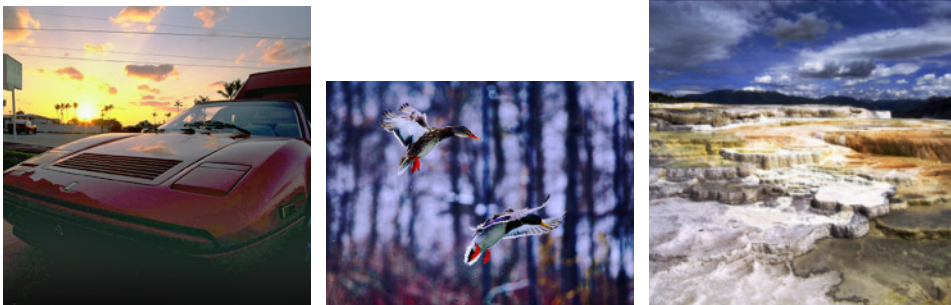


Figure 6: Outputs of the algorithm corresponding to the energy $E_{w,\epsilon}^{\text{log}}$.

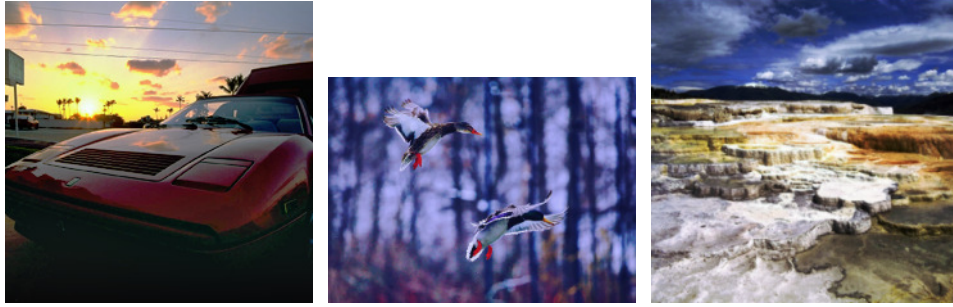


Figure 7: Outputs of the algorithm corresponding to the energy $E_{w,\epsilon}^{-\mathcal{M}}$.

Tests confirm that the considered algorithms indeed enhance contrast, remove color cast and equalize both underexposed and overexposed pictures. As expected, the algorithms derived from energies $E_{w,\epsilon}^{\text{id}}$ and $E_{w,\epsilon}^{-\mathcal{M}}$ have very similar enhancement characteristics, while they differ from the one based on the energy $E_{w,\epsilon}^{\log}$ in the treatment of low and high gray levels. This latter algorithm has greater contrast enhancement properties in dark regions, but the ones corresponding to $E_{w,\epsilon}^{\text{id}}$ and $E_{w,\epsilon}^{-\mathcal{M}}$ perform better in bright areas.

Moreover, regarding the treatment of overexposed images, it is interesting to compare the effect of the ‘pure’ ratio-reset Retinex action (without post-LUT procedures) with the results of the algorithms relative to formulas (26) and (28) on the image shown in fig. 4 (right). In fig. 8 it can be noticed that the result of Retinex is a further overexposed image, while the other two algorithms are able to equalize the image.



Figure 8: *Left.* Action of Retinex (in the RSR implementation [31]). *Center.* Action of the algorithm corresponding to (26). *Right.* Action of the algorithm corresponding to (28).

8.2 Gamma transformation of the basic contrast variable

As just commented, the algorithms based on the energies $E_{w,\epsilon}^{\text{id}}$ and $E_{w,\epsilon}^{-\mathcal{M}}$ have a good behavior in bright areas, but it would be desirable to increase their contrast enhancement ability in dark ones. In order to do that, we have modified the corresponding contrast terms by adding a gamma transformation. Non-linear operations as gamma transformations are also performed by the HVS [4]. Mannos and Sakrison [33] have established that a power

law transformation of the form $I_{\text{in}}(x, y)^\gamma$, being $I_{\text{in}}(x, y)$ the retinal signal and $\gamma \simeq 1/2$, provides good agreement with experimental evidence. Precisely, chosen a $\gamma \in (0, 1)$, we have re-defined the basic contrast variable as $\left(\frac{\min(I(x), I(y))}{\max(I(x), I(y))}\right)^\gamma$. Replacing the basic contrast regularized variable with the gamma-transformed one, we get new energy functionals that still comply with all our assumptions, being the gamma transformation strictly monotone. Explicitly, we redefine the energies $E_{w, \epsilon}^{\text{id}}$ and $E_{w, \epsilon}^{-\mathcal{M}}$ by changing the contrast terms $C_{w, \epsilon}^{\text{id}}$ and $C_{w, \epsilon}^{-\mathcal{M}}$ by the gamma-transformed ones. Using the properties of the logarithmic function, it is easy to see that the consequence of the gamma transformation on $C_w^{\log}(I)$ simply amounts to an overall coefficient change in the corresponding R -function: $R_{\epsilon, \gamma, I^k}^{\log}(x) := \gamma \sum_{y \in \mathcal{I}} w(x, y) s_\epsilon(I(x) - I(y))$, so we will not consider this trivial case in our discussion.

$$C_{w, \epsilon, \gamma}^{\text{id}}(I) = \frac{1}{4\gamma} \sum_{x \in \mathcal{I}} \sum_{y \in \mathcal{I}} w(x, y) \left(\frac{\min_\epsilon(I(x), I(y))}{\max_\epsilon(I(x), I(y))} \right)^\gamma; \quad (36)$$

$$C_{w, \epsilon, \gamma}^{-\mathcal{M}}(I) = -\frac{1}{4\gamma} \sum_{x \in \mathcal{I}} \sum_{y \in \mathcal{I}} w(x, y) \frac{A_\epsilon(I(x)^\gamma - I(y)^\gamma)}{I(x)^\gamma + I(y)^\gamma}. \quad (37)$$

By direct computation, it can be seen that the gradient descent equations corresponding to the gamma-transformed contrast terms still have the general form represented by equation (25), but now with these new R -functions

$$R_{\epsilon, \gamma, I^k}^{\text{id}}(x) := \sum_{y \in \mathcal{I}} w(x, y) \left(\frac{\min_\epsilon(I^k(x), I^k(y))}{\max_\epsilon(I^k(x), I^k(y))} \right)^\gamma s_\epsilon(I^k(x) - I^k(y)), \quad (38)$$

$$R_{\epsilon, \gamma, I^k}^{-\mathcal{M}}(x) := \sum_{y \in \mathcal{I}} w(x, y) \frac{2 (\min_\epsilon(I^k(x), I^k(y))^\gamma \max_\epsilon(I^k(x), I^k(y))^\gamma)}{(\min_\epsilon(I^k(x), I^k(y))^\gamma + \max_\epsilon(I^k(x), I^k(y))^\gamma)^2} s_\epsilon(I^k(x) - I^k(y)). \quad (39)$$

The consequences of the gamma transformation on the analytic form of the r -functions $r_{\epsilon, \gamma}^{\text{id}}(I(x), I(y))$ and $r_{\epsilon, \gamma}^{-\mathcal{M}}(I(x), I(y))$, appearing in $R_{\epsilon, \gamma, I^k}^{\text{id}}(x)$ and $R_{\epsilon, \gamma, I^k}^{-\mathcal{M}}(x)$, respectively, can be appreciated in fig. 9.

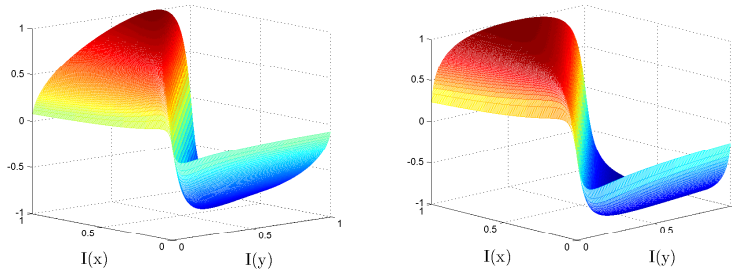


Figure 9: Graphics of $r_{\epsilon, \gamma}^{\text{id}}$ (left) and $r_{\epsilon, \gamma}^{-\mathcal{M}}$ (right) with $\gamma = 1/2$ and $s_\epsilon(z) = \arctan(z/\epsilon) / \arctan(1/\epsilon)$ with $\epsilon = 1/20$.

It can be noticed that the gradient of both surfaces around $(0, 0)$ increased, while they both get slightly smoothed for higher values. The smoothing gets stronger as γ decreases, until the limit $\gamma = 0$, where they reduce to the same surface, precisely: $r_\epsilon^{\log}$. The most suitable value of γ depends on the image content, however our tests showed that an overall good performance with $\gamma = 1/2$.

The pictures in figs. 10 and 11 confirm that the gamma transformed algorithms have greater contrast enhancement properties in dark areas with respect to the original ones, while maintaining good enhancement properties in bright zones.

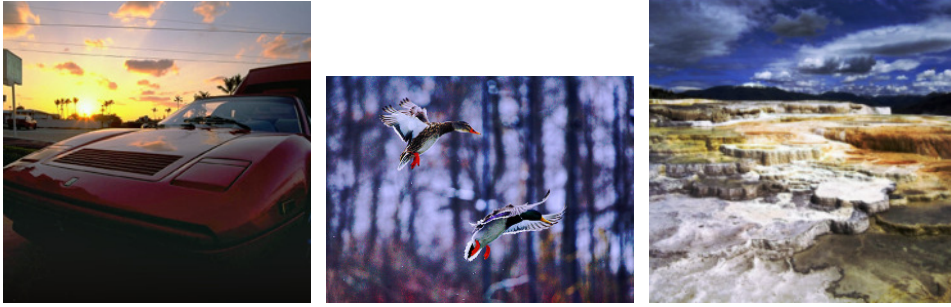


Figure 10: Outputs of the algorithm corresponding to the energy $E_{w,\epsilon}^{\text{id}}$ with gamma transformed contrast variable, $\gamma = 1/2$.

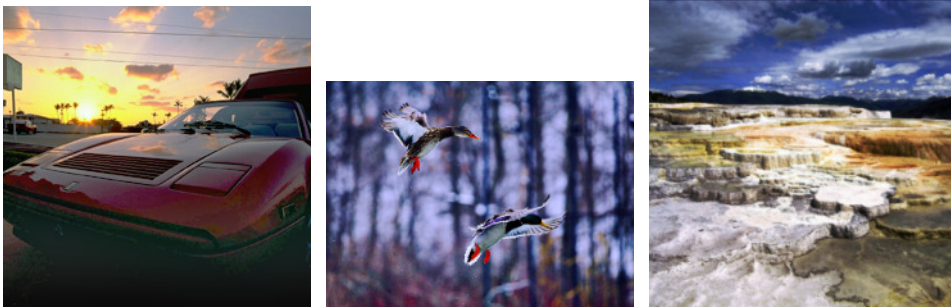


Figure 11: Outputs of the algorithm corresponding to the energy $E_{w,\epsilon}^{-\mathcal{M}}$ with gamma transformed contrast variable, $\gamma = 1/2$.

We present other examples of results. To save space we present the outputs relative only to the algorithm corresponding to the energy $E_{w,\epsilon}^{-\text{id}}$ with gamma transformed contrast variable, $\gamma = 1/2$.

8.3 Tests on cognitive images

We have also tested the ability of our algorithms to reproduce optical illusions. Here we present in Figs. 13 and 14 our results on two classical images: Mach bands and simultaneous contrast, respectively. As can be seen, the results are coherent with the experimental evidences.

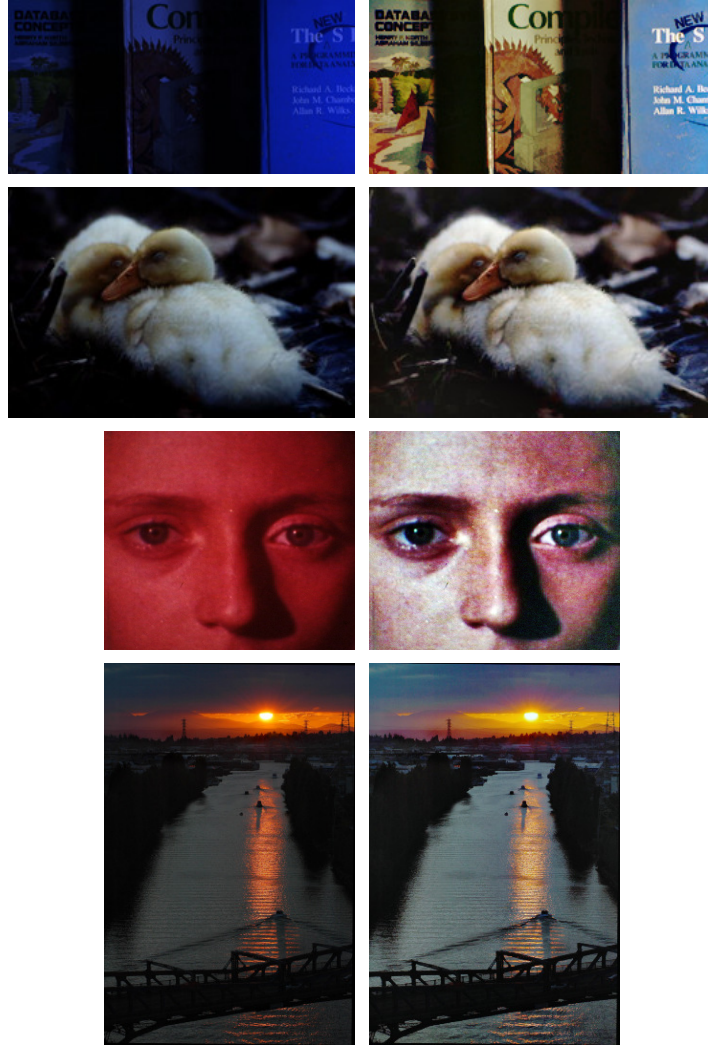


Figure 12: *Left*: original images. *Right*: filtered versions with the algorithm corresponding to the energy $E_{w,\epsilon}^{-\text{id}}$ with gamma transformed contrast variable, $\gamma = 1/2$.

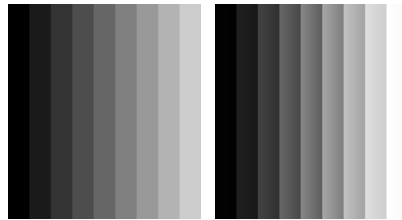


Figure 13: *Left*: original Mach band image. *Right*: filtered version with the algorithm corresponding to the energy $E_{w,\epsilon}^{-\text{id}}$ with gamma transformed contrast variable, $\gamma = 1/2$.

8.4 A pre and post-processing strategy for noise control

Even though the algorithms we are discussing do not *introduce* noise in images, they can *enhance* noise if it is already present. In particular, there can be ‘limit’ situations in which

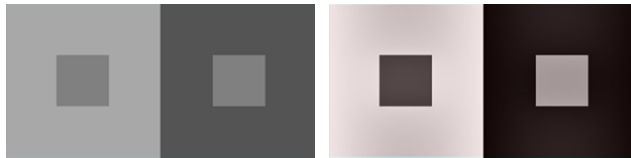


Figure 14: *Left*: original simultaneous contrast image. *Right*: filtered version with the algorithm corresponding to the energy $E_{w,\epsilon}^{-\text{id}}$ with gamma transformed contrast variable, $\gamma = 1/2$.

an additional noise control strategy is needed. This is the case, e.g., for the image shown in fig. 15 (left): the dark area at the bottom right is not completely uniform, but it is filled with noisy pixels that are difficult to distinguish. However, the contrast enhancement produced by our algorithms reveals them, as the unpleasant image in fig. 15 (center) shows.



Figure 15: *Left*: A low key image. *Center*: its filtered version after the application of the algorithm derived by $E_{w,\epsilon}^{\text{id}}$ (the algorithms relative to the other energies produce the same effect). *Right*: noise control provided by the strategy exposed in the text.

To avoid this undesirable effect, we have devised the following strategy. First of all, we apply on the input image an *extrema killer*, or Luc Vincent grain filter, to remove intensity peaks and valleys, see [34]. Secondly, we apply one of our color enhancement algorithms to the pre-filtered images. Finally, we add back the details that were removed in the first step.

Our tests have shown that this procedure works in a satisfactory way for very dark images as the one in fig. 15, but it tends to slightly blur brighter images or pictures with a greater amount of details. For this reason, a user can choose to add or not this strategy to the filtering procedure, depending on the particular need or on image features. Possibly, an automatization of this choice could also be possible: the user could define a threshold limit for global image intensity and variance, in order to filter only those pictures whose intensity and amount of detail (measured, e.g., by variance) are below a certain level.

9 Conclusions and future works

We have presented an investigation about perceptually inspired color enhancement from the point of view of variational techniques. Instead of explicitly devising an algorithm inspired by some Human Visual System features, we have translated the basic phenomenology of

human color vision into mathematical assumptions to be fulfilled by perceptual energy functionals. Operative algorithms for perceptual color correction can then be obtained by minimizing this kind of functionals.

The advantage to analyze perceptual energy functionals is that they can put in evidence in a clearer way the action of color correction in terms of important image features such as, e.g., contrast and dispersion. In fact, our set of assumptions determines a particular class of functionals, whose minimization corresponds to two opponent mechanisms: on one side a local and non-linear contrast enhancement, on the other side a control of dispersion around the middle gray and around the input image data values. The minimization is implemented through an iterative gradient descent technique; the stability of this method and the existence of minima have been proven.

We also pointed out theoretical similarities and differences between these models and existing perceptual color correction models, Retinex and ACE, by identifying, in the class of perceptual functionals, an energy functional that leads to ACE and one that corresponds to a sort of symmetrization of Retinex.

Tests on different images confirm theoretical predictions about the proposed algorithms, which can enhance details and remove color cast without introducing noise. However, already present noise can eventually be amplified. For this reason we also provided a simple and fast pre and post processing procedure that can be integrated in the algorithms to avoid noise amplification in very dark images.

The algorithms depend on four explicit parameters with a clear meaning. Even though there is a set of parameters that works in a satisfactory way for all images of our test set, an image dependent tuning could possibly improve the enhancement characteristics. However, this tuning is beyond the scope of this paper and we consider it as an open problem.

Finally, we believe that the noise control procedure presented in this paper can be improved, e.g. treating the problem from a functional point of view and investigating the relationship between noise and local spatial frequency.

Acknowledgments

R. Palma-Amestoy acknowledges support by Alfa CVFA, AML// 19.0902/97/0666/II-0366-FA. E. Provenzi acknowledges partial support by PRIN-MIUR research project, 2005115173-002. M. Bertalmío and V. Caselles acknowledge partial support by PNPGC project, reference MTM2006-14836, and IP-RACINE Project IST-511316.

Appendix

9.1 The proof of Proposition 5.1

Notice that since $c : (0, 1] \times (0, 1] \rightarrow \mathbb{R}$ is continuous, then $c(I(x), I(y))$ is a bounded function for any $I \in \mathcal{F}_\rho := \{I : \mathfrak{I} \rightarrow [0, 1], I(x) \geq \rho \ \forall x \in \mathfrak{I}\}$. Then the term $C_w(I)$ is bounded in $\mathcal{F}_\rho$. The same can be said of $D_{\alpha, \beta}^\varepsilon(I)$. In particular, $\inf_{I \in \mathcal{F}_\rho} E(I) > -\infty$. Let

$I_k \in \mathcal{F}_\rho$ be a minimizing sequence, i.e., assume that $E(I_k) \rightarrow \inf_{I \in \mathcal{F}_\rho} E(I)$ as $k \rightarrow \infty$. Since I_k is uniformly bounded and we are working in a finite dimensional space, there exists a subsequence, denote it again by I_k , that converges to some $I_\infty \in \mathcal{F}_\rho$. Since both terms $D_{\alpha,\beta}^\varepsilon(I)$ and $C_w(I)$ are continuous in $\mathcal{F}_\rho$, then $E(I_\infty) = \inf_{I \in \mathcal{F}_\rho} E(I)$. The corresponding result for the quadratic dispersion term $D_{\alpha,\beta}^q(I)$ is proved in the same way.

9.2 Examples of nice regularization of the absolute value

Let us first check that the examples of regularization of $A(z)$ given in subsection 5 are nice regularization. Recall that we denote by $O(F(\epsilon))$ any expression satisfying $|O(F(\epsilon))| \leq CF(\epsilon)$ for some $C > 0$ and $\epsilon > 0$ small enough. The examples are:

a) $A_\epsilon(z) = \sqrt{\epsilon^2 + |z|^2} - \epsilon$: in this case $s_\epsilon(z) = \frac{z}{\sqrt{\epsilon^2 + |z|^2}}$. Then $|A_\epsilon(z) - |z|| = \epsilon + |z| - \sqrt{\epsilon^2 + |z|^2} \leq \epsilon$, that is $Q_{1,\epsilon}(z) = O(\epsilon)$. On the other hand, $Q_{2,\epsilon}(z) := A_\epsilon(z) - zs_\epsilon(z) = \frac{\epsilon^2}{\sqrt{\epsilon^2 + |z|^2}} - \epsilon = O(\epsilon)$ as $\epsilon \rightarrow 0$ uniformly in $z \in [-1, 1]$. The other conditions are immediate and we do not give the details.

b) $A_\epsilon(z) = z \frac{\arctan(z/\epsilon)}{\arctan(1/\epsilon)} - \frac{\epsilon}{2 \arctan(1/\epsilon)} \log(1 + \frac{z^2}{\epsilon^2})$: in this case $s_\epsilon(z) = \frac{\arctan(z/\epsilon)}{\arctan(1/\epsilon)}$. Then

$$A_\epsilon(z) = |z| + |z| \frac{(\arctan(|z|/\epsilon) - \arctan(1/\epsilon))}{\arctan(1/\epsilon)} - \frac{\epsilon/2 \log(1 + \frac{z^2}{\epsilon^2})}{\arctan(1/\epsilon)}.$$

Since we have for any $z \in [-1, 1]$,

$$|\arctan(1/\epsilon) - \arctan(|z|/\epsilon)| = \int_{|z|/\epsilon}^{1/\epsilon} \frac{1}{1+s^2} ds \leq \int_{|z|/\epsilon}^{1/\epsilon} \frac{1}{s^2} ds = \epsilon \left(\frac{1}{|z|} - 1 \right),$$

we have

$$Q_{1,\epsilon}(z) = |A_\epsilon(z) - |z|| \leq \frac{\epsilon}{\arctan(1/\epsilon)} + \frac{\epsilon/2 \log(1 + \frac{z^2}{\epsilon^2})}{\arctan(1/\epsilon)} = O(\epsilon \log(1/\epsilon))$$

uniformly in $z \in [-1, 1]$. Finally,

$$Q_{2,\epsilon}(z) = A_\epsilon(z) - zs_\epsilon(z) = -\frac{\epsilon/2 \log(1 + \frac{z^2}{\epsilon^2})}{\arctan(1/\epsilon)} = O(\epsilon \log(1/\epsilon)),$$

uniformly in $z \in [-1, 1]$. Again, the other conditions are immediate and we do not give the details.

9.3 The proof of Proposition 5.2

Before going into the proof of Proposition 5.2, in the following Lemma we compute the first variation of a general functional in presence of symmetries.

Lemma 9.1 Let $w : \mathfrak{I}^2 \rightarrow \mathbb{R}$ be a symmetric function in (x, y) and $F : (0, \infty)^2 \rightarrow \mathbb{R}$ be a differentiable function in its variables (a, b) . Assume that F is symmetric, i.e. $F(a, b) = F(b, a)$ for any $(a, b) \in (0, \infty)^2$. Let $F_1(a, b) = \frac{\partial F}{\partial a}(a, b)$. Then, given

$$E(I) = \iint_{\mathfrak{I}^2} w(x, y) F(I(x), I(y)) dx dy, \quad (40)$$

its first variation can be written as

$$\delta E(I) = 2 \int_{\mathfrak{I}} w(x, y) F_1(I(x), I(y)) dy. \quad (41)$$

Proof. Let $F_2(a, b) = \frac{\partial F}{\partial b}(a, b)$. Since $F(a, b) = F(b, a)$, for all $a, b > 0$, we have

$$F_1(a, b) = F_2(b, a). \quad (42)$$

By definition, the first variation of $E(I)$ in the direction δI is

$$\delta E(I, \delta I) = \iint_{\mathfrak{I}^2} w(x, y) F_1(I(x), I(y)) \delta I(x) dx dy + \iint_{\mathfrak{I}^2} w(x, y) F_2(I(x), I(y)) \delta I(y) dx dy. \quad (43)$$

Interchanging the role of x and y in the second integral of the equation above and using (42) we get

$$\iint_{\mathfrak{I}^2} w(x, y) F_2(I(y), I(x)) \delta I(x) dx dy = \iint_{\mathfrak{I}^2} w(x, y) F_1(I(x), I(y)) \delta I(x) dx dy \quad (44)$$

so that

$$\delta E(I, \delta I) = \int_{\mathfrak{I}} \left(2 \int_{\mathfrak{I}} w(x, y) F_1(I(x), I(y)) \right) \delta I(x) dx \quad (45)$$

and the proposition follows. $\square$

Proof of Proposition 5.2. We assume that $A_\epsilon(z)$ is a nice regularization of the absolute value $A(z) = |z|$, $z \in \mathbb{R}$. We define the regularization $\max_\epsilon(a, b) = \frac{1}{2}(a + b - A_\epsilon(a, b))$ and $\min_\epsilon(a, b) = \frac{1}{2}(a + b + A_\epsilon(a, b))$ for any $a, b \in \mathbb{R}$. Since the integrand in $C_{w, \epsilon}^{\text{id}}(I)$, $C_{w, \epsilon}^{\text{log}}(I)$, $C_{w, \epsilon}^{-\mathcal{M}}(I)$ is differentiable, the result follows by applying Lemma 9.1.

(i) The functional $C_{w, \epsilon}^{\text{id}}(I)$ can be written as (40) taking $F(a, b) = \frac{1}{4} \frac{\min_\epsilon(a, b)}{\max_\epsilon(a, b)}$. Since

$$F_1(a, b) = \frac{1}{8} \frac{A_\epsilon(a - b) - s_\epsilon(a - b)(a + b)}{\max_\epsilon(a, b)^2} = -\frac{1}{4} \frac{b}{\max_\epsilon(a, b)^2} s_\epsilon(a - b) + \frac{1}{8} \frac{Q_{2, \epsilon}(a - b)}{\max_\epsilon(a, b)^2},$$

by Lemma 9.1, we have

$$\begin{aligned} \delta C_{w, \epsilon}^{\text{id}}(I) &= -\frac{1}{2} \sum_{y \in \mathfrak{I}} w(x, y) \left(\frac{I(y) s_\epsilon(I(x) - I(y))}{\max_\epsilon(I(x), I(y))^2} - \frac{1}{2} \frac{Q_{2, \epsilon}(I(x) - I(y))}{\max_\epsilon(I(x), I(y))^2} \right) \\ &= -\frac{1}{2} \sum_{y \in \mathfrak{I}} w(x, y) \frac{I(y)}{\max_\epsilon(I(x), I(y))^2} s_\epsilon(I(x) - I(y)) + S_\epsilon, \end{aligned} \quad (46)$$

where $S_\epsilon = O(Q_{2,\epsilon}(I(x) - I(y)))$, Now, since $A_\epsilon(z) = |z| + Q_{1,\epsilon}(z)$, then $\max_\epsilon(a, b) = \max(a, b) + \frac{1}{2}Q_{1,\epsilon}(z)$, and we may write

$$\frac{1}{\max_\epsilon(a, b)^2} = \frac{1}{\max(a, b)^2} + O(Q_{1,\epsilon}(a, b))$$

where the expression $Q_{1,\epsilon}(a, b)$ is uniform in $a, b \in [\rho, 1]$. Thus, we obtain

$$\delta C_{w,\epsilon}^{\text{id}}(I) = -\frac{1}{2} \sum_{y \in \mathcal{I}} w(x, y) \frac{I(y)}{\max(I(x), I(y))^2} s_\epsilon(I(x) - I(y)) + S'_\epsilon, \quad (47)$$

where $S'_\epsilon = O(Q_{1,\epsilon}(I(x) - I(y)) + Q_{2,\epsilon}(I(x) - I(y)))$.

(ii) The functional $C_{w,\epsilon}^{\log}(I)$ can be written as (40) taking $F(a, b) = \frac{1}{4} \log \frac{\min_\epsilon(a, b)}{\max_\epsilon(a, b)}$. Since

$$\begin{aligned} F_1(a, b) &= \frac{1}{8} \frac{A_\epsilon(a - b) - s_\epsilon(a - b)(a + b)}{\min_\epsilon(a, b) \max_\epsilon(a, b)} \\ &= -\frac{1}{4} \frac{b}{\min_\epsilon(a, b) \max_\epsilon(a, b)} s_\epsilon(a - b) + \frac{1}{8} \frac{Q_{2,\epsilon}(a - b)}{\min_\epsilon(a, b) \max_\epsilon(a, b)}, \end{aligned}$$

by Lemma 9.1, we have

$$\delta C_{w,\epsilon}^{\log}(I) = -\frac{1}{2} \sum_{y \in \mathcal{I}} w(x, y) \left(\frac{I(y) s_\epsilon(I(x) - I(y))}{(I(x) \cdot I(y))_\epsilon} - \frac{1}{2} \frac{Q_{2,\epsilon}(I(x) - I(y))}{(I(x) \cdot I(y))_\epsilon} \right), \quad (48)$$

where $(a \cdot b)_\epsilon := \min_\epsilon(a, b) \max_\epsilon(a, b)$, $a, b \in \mathbb{R}$. Now, since $A_\epsilon(z) = |z| + Q_{1,\epsilon}(z)$, we have

$$\frac{1}{\min_\epsilon(a, b) \max_\epsilon(a, b)} = \frac{1}{\min(a, b) \max(a, b)} + O(Q_{1,\epsilon}(a, b)) = \frac{1}{ab} + O(Q_{1,\epsilon}(a, b))$$

where the expression $Q_{1,\epsilon}(a, b)$ is uniform in $a, b \in [\rho, 1]$. Thus, we obtain

$$\delta C_{w,\epsilon}^{\log}(I) = -\frac{1}{2} \sum_{y \in \mathcal{I}} w(x, y) \frac{1}{I(x)} s_\epsilon(I(x) - I(y)) + S_\epsilon, \quad (49)$$

where $S_\epsilon = O(Q_{1,\epsilon}(I(x) - I(y)) + Q_{2,\epsilon}(I(x) - I(y)))$.

(iii) The functional $C_{w,\epsilon}^{-\mathcal{M}}(I)$ can be written as (40) taking $F(a, b) = -\frac{1}{4} \frac{A_\epsilon(a - b)}{a + b}$. Since

$$F_1(a, b) = -\frac{1}{2} \frac{b}{(a + b)^2} s_\epsilon(a - b) + \frac{1}{4} \frac{Q_{2,\epsilon}(a - b)}{(a + b)^2},$$

by Lemma 9.1, we have

$$\begin{aligned} \delta C_{w,\epsilon}^{-\mathcal{M}}(I) &= -\sum_{y \in \mathcal{I}} w(x, y) \left(\frac{I(y) s_\epsilon(I(x) - I(y))}{(I(x) + I(y))^2} - \frac{1}{2} \frac{Q_\epsilon(I(x) - I(y))}{(I(x) + I(y))^2} \right) \\ &= -\sum_{y \in \mathcal{I}} w(x, y) \frac{I(y)}{(I(x) + I(y))^2} s_\epsilon(I(x) - I(y)) + S_\epsilon, \end{aligned}$$

where $S_\epsilon = O(Q_{2,\epsilon}(I(x) - I(y)))$.

9.4 Remarks on the stability of the algorithm

The general form of semi-implicit gradient descent equation is

$$I^{k+1}(x) = \frac{I^k(x) + \Delta t \left(\frac{\alpha}{2} + \beta I_0(x) + \frac{1}{2} R_{I^k}(x) \right)}{1 + \Delta t(\alpha + \beta)}, \quad (50)$$

being $R_I(x) = \sum_{y \in \mathfrak{I}} w(x, y) r(I(x), I(y))$, for a suitable integrand function r . To fix ideas, we consider the case where

$$R_{\epsilon, I}^{\text{id}}(x) := \sum_{y \in \mathfrak{I}} w(x, y) \frac{\min(I(x), I(y))}{\max(I(x), I(y))} s_\epsilon(I(x) - I(y)). \quad (51)$$

We assume that $I_0 : \mathfrak{I} \rightarrow [\rho, 1]$ where $\rho > 0$ is a minimum value for the initial image. For us $\rho = 1/255$, and this means that we assume that I_0 does not take the value 0. Our purpose is to prove the following statements.

Lemma 9.2 *Assume that $\alpha \geq \frac{1}{1-2\rho} > 0$. Then*

- (i) $\rho \leq I^k(x) \leq 1$ for any $x \in \mathfrak{I}$ and any $k \geq 1$.
- (ii) $\|I^{k+1} - I^k\|_p \leq \frac{1+\Delta t(\frac{1}{\rho}+m_\epsilon)}{1+\Delta t(\alpha+\beta)} \|I^k - I^{k-1}\|_p$ for any $p \in [1, \infty]$ and any $k \geq 1$, where $m_\epsilon := \max_{z \in [-1, 1]} |s'_\epsilon(z)|$. Thus, if $\alpha + \beta > \frac{1}{\rho} + m_\epsilon$, then the iterative scheme given in (50) is convergent to the unique function I^* satisfying

$$\alpha(I^* - \frac{1}{2}) + \beta(I^* - I_0) - \frac{1}{2} R_{\epsilon, I^*}^{\text{id}} = 0. \quad (52)$$

As usual, the $\|\cdot\|_p$ norm of a vector $v = (v_i)_{i=1}^n \in \mathbb{R}^n$, $n \geq 1$, is defined by $\|v\|_p := (\sum_{i=1}^n |v_i|^p)^{1/p}$ if $p \in [1, \infty)$, and $\|v\|_\infty := \max_{i=1, \dots, n} |v_i|$ if $p = \infty$.

Notice that (52) is essentially (and not exactly, due to our regularization of the basic contrast variable) the Euler-Lagrange equation corresponding to the energy $I_{w, \epsilon}^{\text{id}}$. Notice also that $\alpha + \beta > \frac{1}{\rho} + m_\epsilon$ is not a reasonable condition in practice. The reason is that $1/\rho = 255$ and $m_\epsilon \approx 1/\epsilon$ for the particular nice regularization of $A(z)$ given in the examples a), b) given in subsection 5. Then the values of α and/or β are too big and produce a strong attachment to the initial data and/or the value $1/2$; in this case we do not have enough enhancement power. The values of α and β used in practice are much smaller, $\alpha = 255/253$ (in accordance with the assumption of Lemma 9.2) and $\beta = 1$, and the algorithm exhibited stability and convergence in practice.

To prove Lemma 9.2, we need the auxiliary result.

Lemma 9.3 *Let $I, \tilde{I} : \mathfrak{I} \rightarrow [\rho, 1]$. Then*

- (i) $-1 \leq R_{\epsilon, I}^{\text{id}}(x) \leq 1$ for any $x \in \mathfrak{I}$.
- (ii) $\|R_{\epsilon, I}^{\text{id}} - R_{\epsilon, \tilde{I}}^{\text{id}}\|_p \leq 2 \left(\frac{1}{\rho} + m_\epsilon \right) \|I - \tilde{I}\|_p$

Proof. (i) The first statement follows easily from

$$|R_{\epsilon, I}^{\text{id}}(x)| \leq \sum_{y \in \mathcal{I}} w(x, y) \frac{\min(I(x), I(y))}{\max(I(x), I(y))} |s_\epsilon(I(x) - I(y))| \leq \sum_{y \in \mathcal{I}} w(x, y) = 1. \quad (53)$$

(ii) We have

$$|R_{\epsilon, I}^{\text{id}}(x) - R_{\epsilon, \tilde{I}}^{\text{id}}(x)| \leq \sum_{y \in \mathcal{I}} w(x, y) \left| \frac{\min(I(x), I(y))}{\max(I(x), I(y))} s_\epsilon(I(x) - I(y)) - \frac{\min(\tilde{I}(x), \tilde{I}(y))}{\max(\tilde{I}(x), \tilde{I}(y))} s_\epsilon(\tilde{I}(x) - \tilde{I}(y)) \right|. \quad (54)$$

Let us prove that for any $a, b, c, d, \in [\rho, 1]$ we have

$$\left| \frac{\min(a, b)}{\max(a, b)} s_\epsilon(a - b) - \frac{\min(c, d)}{\max(c, d)} s_\epsilon(c - d) \right| \leq \left(\frac{1}{\rho} + m_\epsilon \right) (|a - c| + |b - d|), \quad (55)$$

where $m_\epsilon = \max_{z \in [-1, 1]} |s'_\epsilon(z)|$. Indeed, we have

$$\begin{aligned} \left| \frac{\min(a, b)}{\max(a, b)} s_\epsilon(a - b) - \frac{\min(c, d)}{\max(c, d)} s_\epsilon(c - d) \right| &\leq \left| \frac{\min(a, b)}{\max(a, b)} - \frac{\min(c, d)}{\max(c, d)} \right| |s_\epsilon(c - d)| \\ &\quad + \left| \frac{\min(c, d)}{\max(c, d)} \right| |s_\epsilon(a - b) - s_\epsilon(c - d)|. \end{aligned} \quad (56)$$

Now, we apply the intermediate value theorem to the function $F_\epsilon(z_1, z_2) = \frac{\min_\epsilon(z_1, z_2)}{\max_\epsilon(z_1, z_2)}$, $z_1, z_2 \in [\rho, 1]$ and we obtain

$$|F_\epsilon(a, b) - F_\epsilon(c, d)| \leq \max_{i=1,2} \max_{z_1, z_2 \in [\rho, 1]} \left| \frac{\partial}{\partial z_i} F_\epsilon(z_1, z_2) \right| (|a - c| + |b - d|)$$

Letting $\epsilon \rightarrow 0$ we obtain,

$$\left| \frac{\min(a, b)}{\max(a, b)} - \frac{\min(c, d)}{\max(c, d)} \right| \leq \frac{1}{\rho} (|a - c| + |b - d|). \quad (57)$$

Introducing (57) in (56) and using again the intermediate value theorem with s_ϵ we obtain (55). Now, introducing (55) into (54) we obtain

$$|R_{\epsilon, I}^{\text{id}}(x) - R_{\epsilon, \tilde{I}}^{\text{id}}(x)| \leq \left(\frac{1}{\rho} + m_\epsilon \right) \left(|I(x) - \tilde{I}(x)| + \sum_{y \in \mathcal{I}} w(x, y) |I(y) - \tilde{I}(y)| \right) \quad (58)$$

Since the function $r \in [0, \infty) \rightarrow r^p$, $p \geq 1$, is convex, using Jensen's inequality we have

$$\begin{aligned} \left(\sum_{x \in \mathcal{I}} \left(\sum_{y \in \mathcal{I}} w(x, y) |I(y) - \tilde{I}(y)| \right)^p \right)^{1/p} &\leq \left(\sum_{x \in \mathcal{I}} \sum_{y \in \mathcal{I}} w(x, y) |I(y) - \tilde{I}(y)|^p \right)^{1/p} \\ &\leq \left(\sum_{y \in \mathcal{I}} |I(y) - \tilde{I}(y)|^p \right)^{1/p}. \end{aligned}$$

Using this inequality in (58) we obtain

$$\|R_{\epsilon, I}^{\text{id}} - R_{\epsilon, \tilde{I}}^{\text{id}}\|_p \leq 2 \left(\frac{1}{\rho} + m_\epsilon \right) \|I - \tilde{I}\|_p$$

for any $p \in [1, \infty)$. Letting $p \rightarrow \infty$ we also obtain (ii) for $p = \infty$.

Proof of Lemma 9.2. (i) First, let us check that all iterations $I^k(x) \leq 1$ for any $x \in \mathfrak{I}$ and any $k \geq 1$. This follows easily by induction from (50) since, assuming that $I^k(x) \leq 1$ and using that $\alpha \geq 1$ and Lemma (9.3), (i), we have

$$I^{k+1}(x) \leq \frac{1 + \Delta t \left(\frac{\alpha}{2} + \beta + \frac{1}{2} \right)}{1 + \Delta t(\alpha + \beta)} \leq \frac{1 + \Delta t(\alpha + \beta)}{1 + \Delta t(\alpha + \beta)} = 1.$$

Now assume that $I^k(x) \geq \rho$. To prove that $I^{k+1}(x) \geq \rho$, we use that $R_{\epsilon, I}^{\text{id}}(x) \geq -1$ in (50) and we obtain,

$$I^{k+1}(x) \geq \frac{\rho + \Delta t \left(\frac{\alpha}{2} + \beta\rho - \frac{1}{2} \right)}{1 + \Delta t((\alpha - 1)/2 + \beta)} \geq \rho \frac{1 + \Delta t(\alpha + \beta)}{1 + \Delta t(\alpha + \beta)} = \rho,$$

because we are assuming that $\frac{\alpha}{2} - \frac{1}{2} \geq \alpha\rho$.

(ii) We write

$$I^k(x) = \frac{I^{k-1}(x) + \Delta t \left(\frac{\alpha}{2} + \beta I_0(x) + \frac{1}{2} R_{\epsilon, I^{k-1}}^{\text{id}}(x) \right)}{1 + \Delta t(\alpha + \beta)}, \quad (59)$$

and take the difference of (50) and (59), to obtain

$$\|I^{k+1} - I^k\|_p \leq \frac{\|I^k - I^{k-1}\|_p}{1 + \Delta t(\alpha + \beta)} + \frac{1}{1 + \Delta t(\alpha + \beta)} \frac{1}{2} \|R_{\epsilon, I^k}^{\text{id}} - R_{\epsilon, I^{k-1}}^{\text{id}}\|_p$$

Using Lemma 9.3, (ii), we obtain

$$\|I^{k+1} - I^k\|_p \leq \frac{1 + \Delta t \left(\frac{1}{\rho} + m_\epsilon \right)}{1 + \Delta t(\alpha + \beta)} \|I^k - I^{k-1}\|_p$$

for any $p \in [1, \infty]$ and any $k \geq 1$. The above estimate is contractive when $\frac{1 + \Delta t(\frac{1}{\rho} + m_\epsilon)}{1 + \Delta t(\alpha + \beta)} < 1$, i.e, when $\alpha + \beta > \frac{1}{\rho} + m_\epsilon$. In this case, the sequence I^k converges to a function I^* satisfying

$$I^*(x) = \frac{I^*(x) + \Delta t \left(\frac{\alpha}{2} + \beta I_0(x) + \frac{1}{2} R_{\epsilon, I^*}^{\text{id}}(x) \right)}{1 + \Delta t(\alpha + \beta)},$$

that is, (52) holds.

To prove the uniqueness of solutions of (52), assume that I^* and $\tilde{I}^*$ are two solutions. Then, we have

$$I^* - \tilde{I}^* = \frac{1}{\alpha + \beta} \frac{1}{2} (R_{\epsilon, I^*}^{\text{id}} - R_{\epsilon, \tilde{I}^*}^{\text{id}}).$$

Hence, taking norms and using Lemma 9.3, (ii), we have

$$\|I^* - \tilde{I}^*\| \leq \frac{\frac{1}{\rho} + m_\epsilon}{\alpha + \beta} \|I^* - \tilde{I}^*\|.$$

Since $\alpha + \beta > \frac{1}{\rho} + m_\epsilon$ this is a contradiction, unless $I^* = \tilde{I}^*$.

References

- [1] R. Gregory, *Eye and Brain*. Princeton University Press, 1997.
- [2] D. Hubel, *Eye, Brain, and Vision*. Scientific American Library, 1995.
- [3] S. Palmer, *Vision Science: Photons to Phenomenology*. The MIT Press, 1999.
- [4] W. Pratt, *Digital Image Processing*. J. Wiley & Sons, 2007.
- [5] B. Wandell, *Foundation of vision*. Sinauer Associates, 1995.
- [6] G. Wyszecky and W. S. Stiles, *Color science: Concepts and methods, quantitative data and formulas*. John Wiley & Sons, 2000.
- [7] S. Zeki, *Vision of the brain*. Blackwell Publishing Incorporated, 1993.
- [8] E. Land and J. McCann, “Lightness and Retinex theory,” *Journal of the Optical Society of America*, vol. 61, no. 1, pp. 1–11, Jan. 1971.
- [9] E. Land, “The Retinex theory of color vision,” *Scientific American*, vol. 237, pp. 108–128, 1977.
- [10] —, “Recent advances in retinex theory and some implications for cortical computations: Color vision and the natural image,” in *Proc. Nat. Acad. Sci. U.S.A.*, vol. 80, 1983, pp. 5163–5169.
- [11] S. Zeki and L. Marini, “Three cortical stages of colour processing in the human brain,” *Brain*, vol. 121, pp. 1669–1685, 1998.
- [12] J. J. McCann and et al., “Special session on Retinex at 40,” *Journal of Electronic Imaging*, vol. 13, no. 1, pp. 6–145, January 2004.
- [13] A. Rizzi, C. Gatta, and D. Marini, “A new algorithm for unsupervised global and local color correction,” *Pattern Recognition Letters*, vol. 24, pp. 1663–1677, 2003.
- [14] G. West, “Color perception and the limits of color constancy,” *Journal of mathematical biology*, vol. 8, pp. 47–53, 1979.

- [15] B. Funt, K. Barnard, and L. Martin, “Is colour constancy good enough?” in *5th European Conference on Computer Vision*, 1998, pp. 445–459.
- [16] M. Ebner, *Color constancy*. Wiley, 2007.
- [17] A. Hurlbert, “Formal connections between lightness algorithms,” *Journ. Opt. Soc. Am. A*, vol. 3, pp. 1684–1693, 1986.
- [18] R. Gonzales and R. Woods, *Digital image processing*. Prentice Hall, 2002.
- [19] O. Creutzfeld, B. Lange-Malecki, and K. Wortmann, “Darkness induction, retinex and cooperative mechanisms in vision,” *Exp. Brain Res.*, vol. 67, pp. 270–283, 1987.
- [20] O. Creutzfeld, B. Lange-Malecki, and E. Dreyer, “Darkness induction, retinex and cooperative mechanisms in vision,” *Jour. Opt. Soc. Am. A*, vol. 7, pp. 1644–1653, 1990.
- [21] L. Hurvich and D. Jameson, “Theory of brightness and color contrast in human vision,” *Vision Research*, vol. 4, pp. 135–154, 1990.
- [22] Q. Zaidi, *Color and brightness induction: from Mach bands to three-dimensional configurations*. Cambridge University Press, New York, 1999.
- [23] R. Shapley and C. Enroth-Cugell, *Visual adaptation and retinal gain controls*, 1984, vol. 3, pp. 263–346.
- [24] G. Buchsbaum, “A spatial processor model for object colour perception,” *Journal of the Franklin Institute*, vol. 310, pp. 337–350, 1980.
- [25] S. Usui and S. Nakauchi, “A neurocomputational model for color constancy,” in *In C. Dickinson, I. Murray and D. Carden: John Dalton’s Colour Vision Lagacy. Selected Proceedings of the International Conference*. London, UK: Taylor and Francis, 1997, pp. 475–482.
- [26] A. Michelson, *Studies in Optics*. Chicago University Press, 1927.
- [27] L. Ambrosio, N. Gigli, and G. Savaré, *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics, Birkhauser, 2005.
- [28] G. Sapiro and V. Caselles, “Histogram modification via differential equations,” *Jour. of Diff. Eq.*, vol. 135, pp. 238–266, 1997.
- [29] M. Bertalmío, V. Caselles, E. Provenzi, and A. Rizzi, “Perceptual color correction through variational techniques,” *IEEE Trans. on Image Proc.*, vol. 16, pp. 1058–1072, 2007.

- [30] E. Provenzi, L. De Carli, A. Rizzi, and D. Marini, “Mathematical definition and analysis of the Retinex algorithm,” *Journal of the Optical Society of America A*, vol. 22, no. 12, pp. 2613–2621, Dec. 2005.
- [31] E. Provenzi, M. Fierro, A. Rizzi, L. De Carli, D. Gadia, and D. Marini, “Random spray Retinex: A new Retinex implementation to investigate the local properties of the model,” *IEEE Transactions on Image Processing*, vol. 16, pp. 162–171, Jan. 2007.
- [32] E. Provenzi, C. Gatta, M. Fierro, and A. Rizzi, “Spatially variant white patch and gray world method for color image enhancement driven by local contrast,” *regularly accepted by IEEE Transactions on Pattern Analysis and Machine Intelligence, publication scheduled for 2008*.
- [33] J. L. Mannos and D. J. Sakrison, “The effects of a visual fidelity criterion on the encoding of images,” *IEEE Trans. Information Theory, IT-20*, vol. 4, pp. 525–536, 1974.
- [34] P. Soille, *Morphological image analysis, Principles and Applications*. Springer-Verlag, 1999.