



Published in Image Processing On Line on 2015-07-29.
Submitted on 2015-03-02, accepted on 2015-06-01.
ISSN 2105-1232 © 2015 IPOL & the authors CC-BY-NC-SA
This article is available online with supplementary materials,
software, datasets and online demo at
<https://doi.org/10.5201/ipol.2015.131>

An Algorithmic Analysis of Variational Models for Perceptual Local Contrast Enhancement

S. Ferradans¹, R. Palma-Amestoy², E. Provenzi³

¹ ENS, CNRS, PSL, Paris, France (sira.ferradans@ens.fr)

² Woodtech Measurement Systems, Santiago, Chile (ropalma@ing.uchile.cl)

³ Laboratoire MAP5 (UMR CNRS 8145), Université Paris Descartes, Paris, France
(edoardo.provenzi@parisdescartes.fr)

Communicated by Jose-Luis Lisani

Demo edited by Jose-Luis Lisani

Abstract

Color cast cancellation and local contrast enhancement are very important problems in computer vision. In this paper we review the algorithm proposed by Palma-Amestoy et al. [A perceptually inspired variational framework for color enhancement, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21 (2009), pp. 458–474], present results and evaluate the impact of a change in the parameters.

Source Code

The code implements the algorithm presented by Palma-Amestoy et al. [7], with a pre-computation of the polynomials. These polynomials are inputted from text files which are also provided.

The reviewed source code and documentation for this algorithm are available from [the web page of this article](#)¹. Compilation and usage instruction are included in the `README.txt` file of the archive.

Supplementary Material

As supplementary material, a reference dataset is provided, and the Matlab code to generate the polynomial approximation, although it is not necessary to execute it to test the algorithm.

Keywords: color constancy; contrast enhancement; Weber-Fechner contrast; visual adaptation; human visual system.

¹<https://doi.org/10.5201/ipol.2015.131>

1 Introduction

In this paper, we will consider the problem of local contrast enhancement inspired by human perception and its consequences on color cast removal. More specifically, given any digital image, acquired under general illumination conditions, we will try to modify its chromatic attributes so that the resulting image is as close as possible to the human perception of the scene photographed in the picture.

This kind of algorithms are of great importance in different fields that need an illumination-independent representation of the scene such as object recognition, calibration, image segmentation or tracking. They have also proven necessary in computational photography and flickering reduction, to quote but two applications.

In [7] the authors built a general variational framework to perform perceptual color correction. Their work is based on the selection of energy functionals through suitable constraints related to phenomenological properties of the Human Visual System (HVS).

This variational framework allows to embed existing perceptual color correction algorithms and clarify more easily the main similarities and differences among them. The usefulness of this result relies in the fact that the direct equations of these algorithms are very distant from each other, but not the energies they minimize.

The plan of this paper is the following: in Section 2, we introduce the perceptual laws and the mathematical model behind the algorithm. In Section 3, we detail the numerical scheme and present the final algorithm. Finally, in Section 4, we present some results and how they vary when changing the default parameters.

2 A General Variational Framework for Perceptual Color Correction

In this section, we will introduce the basic results of paper [7]. First of all, we will introduce and discuss the meaning of an important result on the variational interpretation of histogram equalization (HE), which is strongly related to the framework proposed in [7]. Then, we will briefly recall four basic HVS features, and, finally, we will show how to modify the energy function of HE in order to take into account these HVS characteristics, which is one of the main results of [7].

Before going through this description, we introduce the notation that will be used in the rest of the paper. Let us denote with $\Omega \subset \mathbb{R}^2$ the *spatial domain* of a digital image, with $|\Omega|$ its area and with $x \equiv (x_1, x_2)$ and $y \equiv (y_1, y_2)$ the coordinates of two arbitrary pixels in Ω . We will always consider a *normalized dynamic range* in $[0, 1]$, so that a *RGB image function* will be denoted with

$$\begin{aligned} \vec{I}: \Omega &\longrightarrow [0, 1] \times [0, 1] \times [0, 1] \\ x &\mapsto (I_R(x), I_G(x), I_B(x)), \end{aligned}$$

where each scalar component $I_c(x)$ defines the intensity level of the pixel $x \in \Omega$ in the red, green and blue channels, respectively.

We stress that we will perform every computation on the scalar components of the image, thus *treating each chromatic component separately*. Therefore, we will avoid the subscript c and write simply $I(x)$ to denote the intensity of the pixel x in a given chromatic channel.

2.1 Variational Interpretation of Histogram Equalization

One of the main inspirations of paper [7] was the following result about the variational interpretation of histogram equalization [11].

Theorem 1: Histogram equalization can be implemented by minimizing the following energy

$$E_{\text{hist eq}}(I) \equiv 2 \int_{\Omega} \left(I(x) - \frac{1}{2} \right)^2 dx - \frac{1}{|\Omega|} \iint_{\Omega^2} |I(x) - I(y)| dx dy. \quad (1)$$

Let us write the HE energy functional as follows: $E_{\text{hist eq}}(I) = D_{\frac{1}{2}}(I) - C(I)$, where the two functional terms $D_{\frac{1}{2}}(I)$ and $C(I)$ are

$$D_{\frac{1}{2}}(I) \equiv 2 \int_{\Omega} \left(I(x) - \frac{1}{2} \right)^2 dx, \quad (2)$$

and

$$C(I) \equiv \frac{1}{|\Omega|} \iint_{\Omega^2} |I(x) - I(y)| dx dy. \quad (3)$$

The minimization of $E_{\text{hist eq}}(I) = D_{\frac{1}{2}}(I) - C(I)$ is achieved through the *minimization of $D(I)$* and *maximization of $C(I)$* .

Dispersion term. $D_{\frac{1}{2}}(I)$ is called *global quadratic dispersion* term around the middle gray level $1/2$ and it is minimized when $I(x) \equiv 1/2$ for all $x \in \Omega$, i.e. the minimization of this term tends to turn I into a uniform gray image.

Contrast term. $C(I)$ is called *global contrast* term and its maximization corresponds to the intensification of the global contrast of the image I , measured through the absolute differences $|I(x) - I(y)|$.

Optimization. The main result of paper [11] is that, if $I^* = \operatorname{argmin}_I E_{\text{hist eq}}(I)$, then I^* has equalized histogram, i.e. all the intensity levels have the same occurrence probability in the image. In [11], the authors prove this result by showing that, if we denote by H the cumulative histogram, then the image function I^* satisfies the equation $H(I^*(x)) = I^*(x)$, for all $x \in \Omega$, which implies the flatness of the histogram of I^* , see e.g. [4].

The minimization of $E_{\text{hist eq}}(I)$ can be achieved through gradient descent. With respect to this, another important result of [11] is that, if I_0 represents any input image, then the initial value problem

$$\begin{cases} \partial_t I = -\delta E_{\text{hist eq}}(I) \\ I(0) = I_0 \end{cases}, \quad (4)$$

has only one solution, where t is the evolution parameter of the iterative gradient descent scheme and the symbol δ represents the first variation of the functional.

Interpretation. The argmin of $E_{\text{hist eq}}(I)$ is the image corresponding to the optimal balance between *two opposite effects*: on one side, the minimization of $D_{\frac{1}{2}}(I)$ tends to *set all the levels to the average gray $1/2$* but, on the other side, the minimization of $-C(I)$ tends to *spread the intensity levels apart, as far as possible from each other*.

So, the intrinsic meaning of the variational description of histogram equalization is that *the equilibrium among two conflicting actions, control of the intensity dispersion around the middle gray level and contrast enhancement, induces histogram equalization*.

One practical consequence of this result is that, applying for example the gradient descent technique to minimize $E_{\text{hist eq}}(I)$, one can stop the minimization process before reaching the complete equalization, thus realizing a *partial equalization* that can nonetheless be useful to avoid the typical over-enhancement of low-key images (see [4] for more details).

However, for the purposes of color image processing, the most important consequences of Theorem 1 are theoretical: in fact, as we will see in the next section, we can modify the functional $E_{\text{hist eq}}(I)$ in such a way that the basic principle of histogram equalization, i.e. the balance between dispersion control and contrast enhancement, is preserved but we can change the analytical form of the terms $D_{\frac{1}{2}}(I)$ and $C(I)$ taking inspiration from human visual perceptual features. The argmin image I^* of the modified functional is a color-corrected version of the original image driven by perceptual properties of the HVS.

2.2 A Basic Set of Four Basic HVS Properties

The HVS is an extraordinarily complex system, nowadays still far from being fully understood. As it would be impossible to provide an even approximate description, here we will just recall the *four basic phenomenological mechanisms of color vision* that will be used to select perceptual functionals. As a general reference for the following subsection, the reader can refer to the very complete book [13].

2.2.1 Adaptation to the Average Luminance Level

The absorption of a photon by one of the three L, M, S cones activates a photo-chemical transduction process which transforms the electromagnetic energy of the photon to a difference of electric potential of the photoreceptor's membrane.

The empirical law that describes this transduction, see e.g. [12], is known as *Michaelis-Menten's equation* (or *Naka-Rushton's equation* when the coefficient n is unitary)

$$r(L) = \frac{\Delta V}{\Delta V_{\max}} = \frac{L^n}{L^n + L_S^n}, \quad (5)$$

where ΔV_{\max} is the highest difference of potential that can be generated, n is a constant (measured as 0.74 for the rhesus monkey), L is the luminance of the viewed object and L_S is the luminance at which the photoreceptor response is half maximal, called the *semisaturation level*, and which is usually associated with the *level of adaptation*. Each type of cone L, M, S is most sensitive in a particular waveband and the semisaturation constant depends on the amount of light in the particular waveband that is absorbed, not on the global luminance of the light source.

Equation 5 has a sigmoid-like behavior that can be interpreted as follows: the photoreceptors response has an intensity *threshold*, below which their output is null, then they *respond fairly linearly* and, finally, the response *saturates* as we approach large values of luminance.

The 'center' of this behavior is the semisaturation value L_S , which is mapped to 1/2, in this sense the retinal cells 'adapt' to each scene average luminance, allowing us to perceive modulations of light intensity around L_S . This adaptive capacity of retinal cells is of paramount importance for human vision, without it we simply would not be able to see when passing from a dim light environment to a very intensely illuminated one, and viceversa.

2.2.2 Local Contrast Enhancement

The HVS operates a local contrast enhancement, with the primary (but not the only) effect of improving edge perception, as proven by visual effects as simultaneous contrast or Mach bands, see e.g. [4]. Edges contain the most important part of the visual information, so it is not surprising that the HVS has developed a way to improve their perception. The most important feature that must be underlined here is that edge enhancement is *local*, i.e. it depends on the local distribution of light intensity around each point.

2.2.3 Color Constancy

If light adaptation is the HVS ability to adapt to different light *intensities*, *color constancy* may be described at first glance as the ability to adapt to different light *spectral content*, in order to perceive colors as constant as possible to changes in illumination. This feature is so intrinsically hardwired in our visual system that we often take it for granted. However, whenever we take two pictures of the same scene with two different illuminants, we see that we must perform a very careful white balance to obtain two images that produce similar color sensations. The psycho-physiological details that allow color constancy are beyond the scopes of this work, for more details see e.g. [5].

Color constancy will be essential for the selection of the analytical form of the contrast enhancement term of perceptual functionals.

2.2.4 Weber's Law

Weber's psychophysical experiments to test contrast perception consist in what follows: a dark-adapted human observer is put in a dim room in front of a white screen on which a narrow beam of light is thrown in the center of the visual field. The luminous intensity I of the beam is increased very slowly and the observer is asked to tell whether he/she can perceive an intensity change. The least perceptible intensity change ΔI is generally called *JND* for *Just Noticeable Difference*.

Weber found that the JND increases proportionally with the luminous intensity, i.e. $\Delta I = K \cdot I$, or, as it is in general reported

$$\frac{\Delta I}{I} \simeq K. \quad (6)$$

Equation (6) is called *Weber's law* and K Weber's constant. Weber's law says that, as we increase the background light I , the difference ΔI must increase proportionally in order to be able to appreciate $I + \Delta I$ as different from I .

2.3 Variational Perceptually-Inspired Color Correction of Digital Images

In [7] the perceptual features recalled in Subsection 2.2 have been used to modify the histogram equalization functional described in Subsection 2.1 as reported in the next theorem.

Theorem 2: The only class of energy functionals complying with all the four HVS features discussed in Section 2.2 is the following

$$E_{\mu, I_0, w, \varphi}(I) = D_{\mu, I_0}(I) + C_{w, \varphi}(I), \quad (7)$$

where

$$D_{\mu, I_0}(I) = \int_{\Omega} \left[\alpha \left(\mu \log \frac{\mu}{I(x)} - (\mu - I(x)) \right) + \beta \left(I_0(x) \log \frac{I_0(x)}{I(x)} - (I_0(x) - I(x)) \right) \right] dx, \quad (8)$$

and

$$C_{w, \varphi}(I) = \iint_{\Omega^2} w(x, y) \varphi \left(\frac{\min\{I(x), I(y)\}}{\max\{I(x), I(y)\}} \right) dx dy, \quad (9)$$

where:

- I_0 is a given channel of the original image function;
- μ is the average value of I_0 ;

- $\alpha, \beta > 0$ are real coefficients which control the attachment to μ and to the original image function values I_0 , respectively;
- $w : \Omega \times \Omega \rightarrow \mathbb{R}^+$ is a spatial kernel, which depends only on the Euclidean distance $\|x - y\|$ between two generic pixels $x, y \in \Omega$ and it is monotonically decreasing with the distance itself, i.e. $w(x, y) := g(\|x - y\|)$ where $g : \mathbb{R}^+ \rightarrow \mathbb{R}$ is monotonically decreasing;
- $\varphi : [0, 1] \rightarrow \mathbb{R}$ is a monotonically increasing differentiable function.

Contrast Term and Color Constancy. To understand the motivation for the analytical shape of the contrast term $C_{w,\varphi}(I)$ let us consider the basic image formation model, i.e. $I(x) = \rho(x) \cdot \lambda$, where $\rho(x)$ represents the reflectance of a point x and λ represents the illuminant (supposed to be constant all over the scene). Since $C_{w,\varphi}(I)$ is written in terms of a ratio, it is evident that it is independent from illuminant changes, thus it is coherent with the color constancy feature. Moreover, as proven in [7], this analytical form is the only one in which color constancy and Weber's law can be combined.

Function φ represents a degree of freedom and it is chosen to be monotonically increasing to prevent contrast reversion between pixels.

Contrast Term and Local Contrast Enhancement. Finally, notice that the minimization of $C_{w,\varphi}(I)$ indeed induces a contrast enhancement. In fact, the ratio $\min\{I(x), I(y)\} / \max\{I(x), I(y)\}$ is minimized when the lowest value between two pixels is decreased and the highest is increased, which, of course, corresponds to an intensification of contrast. The enhancement is spatially local due to the presence of the weighting function w . Typically w is a Gaussian kernel with center in x , its standard deviation σ can be set by the user to increase or decrease the locality of contrast enhancement. Small values of σ push the effect towards *sharpening*, large values of σ instead push towards a global enhancement.

Dispersion Term. The choice of the dispersion term has been guided by dimensional coherence with the contrast term, which has dimension 0 with respect to I . The easiest meaningful candidate is the *entropy functional* D_{μ, I_0} , whose minimization produces a reduction of entropy, i.e. disorder, around the average value μ (which can be different in each chromatic channel) and around the original image I_0 . This last attachment is introduced to avoid an excessive departure from the original intensity values and can be modulated via the coefficients ratio α/β . To understand why D_{μ, I_0} complies with the tasks of a dispersion term, we note that the abstract analytical form of the integrand can be written as $f(s) = a \log \frac{a}{s} - (a - s)$ [1], where $s \in (0, 1]$ is a generic normalized intensity level and $a > 0$ can represent either μ or $I_0(x)$. The first derivative of $f(s)$ is $\frac{df}{ds}(s) = 1 - \frac{a}{s}$, so $\frac{df}{ds}(s) = 0$ if and only if $s = a$. Moreover, the second derivative is $\frac{d^2f}{ds^2}(s) = \frac{a}{s^2} > 0, \forall s$. This implies that the value $s = a$ is a global minimum for $f(s)$. Hence, *minimizing D_{μ, I_0} corresponds to an attachment to μ and around the original data $I_0(x)$.*

2.4 Relationship with Existing Perceptually-Inspired Color Correction Models

If we compare $E_{\text{hist eq}}$ with $E_{\mu, I_0, w, \varphi}(I)$, we can infer that *a perceptually-inspired color correction can be interpreted as a suitable local and non linear version of histogram equalization*. Moreover, within this framework, it is possible to link two well-known models for perceptually-inspired color correction ACE [10] and Retinex [6]. As previously said, this result is quite remarkable, since the original equations of ACE and Retinex seem to be very distant from each other, see [8] and [9] for

more details. This can be done simply by selecting two different analytical expressions of the function φ . In [3] and [2], respectively, it has been proved that:

- If $\varphi \equiv \log$, then the minimization of $E_{\mu, I_0, w, \log}(I)$ leads to the so-called ACE algorithm formulated in [10];
- If $\varphi \equiv \text{id}$, then the minimization of $E_{\mu, I_0, w, \text{id}}(I)$ leads to the continuous symmetrized version of the original Retinex algorithm [6].

In the test section, apart from the identity and the logarithmic function, we will also consider the following choice of $\varphi \equiv -\mathcal{M}$, where $\mathcal{M} \left(\frac{\min(I(x), I(y))}{\max(I(x), I(y))} \right) := \frac{1 - \frac{\min(I(x), I(y))}{\max(I(x), I(y))}}{1 + \frac{\min(I(x), I(y))}{\max(I(x), I(y))}} \equiv \frac{|I(x) - I(y)|}{I(x) + I(y)}$, which corresponds to a novel algorithm.

3 Algorithm

3.1 Stability of the Numerical Scheme for the Minimization of Perceptual Functionals

The Euler-Lagrange equations of $E_{\mu, I_0, w, \varphi}(I)$, i.e. $\delta E_{\mu, I_0, w, \varphi}(I) = 0$, are integral equations that cannot be solved analytically. Moreover, the perceptual functionals are not convex, so we cannot easily find a global minimum. However, in [7] it has been proved that, if we substitute the null values of the input image with $1/255$ (the smallest non-zero value in the normalized dynamic range), and we consider $\alpha, \beta > 1$, then the gradient descent scheme with respect to $\log I$, i.e.

$$\partial_t \log I = -\delta E_{\mu, I_0, w, \varphi}(I), \tag{10}$$

converges to a unique image function, that we can consider the perceptually-inspired color-corrected one. The gradient descent is written in terms of $\log I$ because the logarithmic derivative, i.e. $\partial_t \log I = \frac{1}{I} \partial_t I$ allows re-writing the gradient descent equation as follows

$$\partial_t I = -I \cdot \delta E_{\mu, I_0, w, \varphi}(I), \tag{11}$$

the right-hand side of this equation has dimension 0 (which is the correct perceptual dimension), since $\delta E_{\mu, I_0, w, \varphi}(I)$ has dimension² -1 in terms of I . Moreover, notice that, as proved in [1], the effect of taking the logarithmic image is simply to change the speed of convergence of the gradient descent scheme and not its final result.

Discretization. Defining a finite evolution step Δt and setting $I_{k+1}(x) = I_{k\Delta t}(x)$ for $k = 0, 1, 2, \dots$, from (11) we obtain the following minimization scheme

$$I_{k+1}(x) = \frac{I_k(x) + \Delta t (\alpha\mu + \beta I_0(x) + \frac{1}{2} \mathcal{C}_{w,k}^\varphi(x))}{1 + \Delta t(\alpha + \beta)}, \tag{12}$$

where

$$\mathcal{C}_{w,k}^\varphi(x) = \sum_{y \in \Omega} w(\|x - y\|) \varphi \left(\frac{\min\{I_k(x), I_k(y)\}}{\max\{I_k(x), I_k(y)\}} \right). \tag{13}$$

²In general, the first variation of a homogeneous functional of degree n is a homogeneous functional of degree $n - 1$, in our case $E_{\mu, I_0, w, \varphi}(I)$ has degree 0, so its first variation has degree -1.

3.2 Complexity Reduction of the Contrast Term

The computational complexity of these algorithms is $\mathcal{O}(N^2)$, N being the number of pixels in the image. This high complexity is due to the contrast term: its first variation generates the integral term $\mathcal{C}_{w,k}^\varphi(x)$ in the Euler-Lagrange equations. We present now the approximation technique proposed in [7] to reduce the computational complexity to $\mathcal{O}(N \log N)$.

The basis of the complexity reduction technique is to write \mathcal{C}_w^φ as an expression that separates the dependence between $I(x)$ and $I(y)$. Let us drop the dependence on the iteration k in this section for the sake of simplicity. We rewrite the generic contrast term in the following way

$$\mathcal{C}_w^\varphi(x) = f(I(x)) \sum_{y \in \mathcal{I}} w(\|x - y\|) g(I(y)). \quad (14)$$

By definition, the sum in the previous formula is a convolution, thus we just need to find the functions f and g . Let p be a generic polynomial of order n on the variables $I(x), I(y)$

$$p(I(x), I(y)) = \sum_{j=0}^n \sum_{m=0}^j p_{j-m,m} I(x)^{j-m} I(y)^m = \sum_{j=0}^n f_j(I(x)) I(y)^j, \quad (15)$$

setting $g(I(y)) = I^j(y)$, we can express the contrast term as

$$\tilde{\mathcal{C}}_w^{[\varphi,n]}(x) = \sum_{j=0}^n f_j(I(x)) (w * I^j)(x), \quad (16)$$

where $f_j(I(x)) = \sum_{m=0}^{n-j} p_{m,j} I^m(x)$, and $*$ is the convolution operator. We compute the coefficients f_j of the polynomial of order n by solving $\tilde{p}^{[\varphi,n]}(I(x), I(y)) = \operatorname{argmin}_p \|p - \varphi\|_2$, i.e. $\tilde{p}^{[\varphi,n]}$ is the polynomial of order n with *minimal quadratic distance* with respect to the function φ .

Numerical Algorithm. Note that the coefficients f_j depend on the intensity of the pixel thus they can be precomputed for a meaningful discrete number of values. See Algorithm 1 for more details. The degree n of the polynomial is a parameter that controls the precision of the polynomial approximation. For all numerical tests, it was set to $n = 9$ which amounts to a number of coefficients $C = 55$.

3.3 Final Algorithm

The numerical scheme (12) accelerated with the complexity reduction approach is detailed in Algorithm 2. In practice, this algorithm is applied to each color component independently.

Restrictions on the parameters α, β . Regarding the parameters $\Delta t, \alpha$ and β , it is important to analyze their restrictions to guarantee the algorithm convergence. Let us set $\Delta t \in [0, 1]$ and $\mu = \frac{1}{2}$. The basic condition for convergence is to keep $I_{k+1}(x) \in [0, 1]$ at each step, which can be done by analyzing the boundary intensity values of Equation (12). The minimal value of the pixel $I_{k+1}(x)$ is obtained when all the equation components are minimal

$$I_{k+1}(x) = \frac{\Delta t \left(\frac{1}{2}\alpha - \frac{1}{2}\right)}{1 + \Delta t(\alpha + \beta)} \geq 0, \quad (17)$$

which leads to the condition $\alpha \geq 1$, since $\beta \geq 0$. The analysis on the maximal value $I_{k+1}(x) \leq 1$ confirms this condition on α .

Algorithm 1: Polynomial coefficients computation

Input: Type of contrast $c = \text{id}, \log, -\mathcal{M}$.

Output: $\tilde{p}^{[\varphi, n]} = \{\tilde{p}_i^c\}_{i=0}^C$, where C is the number of coefficients of the polynomial of degree $n = 9$.

1. **Range** $x = \{\frac{1}{255}, \frac{2}{255}, \dots, \frac{255}{255}\}$ and $y = \{\frac{1}{255}, \frac{2}{255}, \dots, \frac{255}{255}\}$
 2. **Approximation of the sign function** $s_\epsilon(z) = \frac{\arctan(z/\epsilon)}{\arctan(1/\epsilon)}$, where $\epsilon = 1/20$.
 3. **Set contrast function** φ
 - if $c = \text{id}$ then $\varphi^{\text{id}}(x, y) = \frac{xy}{\max_\epsilon(x, y)} s_\epsilon(x - y)$
 - if $c = \log$ then $\varphi^{\log}(x, y) = s_\epsilon(x - y)$
 - if $c = -\mathcal{M}$ then $\varphi^{-\mathcal{M}}(x, y) = \frac{2xy}{(x+y)^2} s_\epsilon(x - y)$
 4. **Fit coefficients** Solve the following least squares problem:

$$\tilde{p}^{[\varphi^c, n]} = \arg \min_p \left\| \sum_{j=0}^n \sum_{m=0}^j p_{j-m, m} x^{j-l} y^l - \varphi^c(x, y) \right\|_2.$$
-

Algorithm 2: Contrast enhancement

Input:

- $I_0(x) \in [0, 1]$: input image, where $x \in \Omega$ is the number of pixels.
- c : type of contrast $\text{id}, \log, -\mathcal{M}$
- σ : weight parameter,
- $\alpha \geq 1, \beta \geq 0$: importance of the attachment to $\frac{1}{2}$ and I_0 , respectively.

Output: I^k

1. **Compute Polynomials:** Find $\tilde{p}^{[\varphi^c, n]}$ with Algorithm 1 with c as the parameter.
 2. **Weighting function:** Precompute $w_\sigma(x, y)$ with equations (18) or (19)
 3. **Set constants:** $\Delta t = 0.02$, $\mu = \frac{1}{2}$ and $\text{proj}_{[a, b]}(x) = \max(\min(x, b), a)$
 4. **Optimization:** Iterate until convergence: while $\|I^{k+1} - I^k\|^2 > \epsilon$
 - Precompute: $W_k^j = w * I_k^j$, for $j = \{1, \dots, n\}$
 - For every pixel $x \in \Omega$:
 - $\tilde{\mathcal{C}}_{w, k}^{[\varphi^c, n]}(x) = \text{proj}_{[-1, 1]} \left(\sum_{j=0}^n \sum_{m=0}^{n-j} \tilde{p}_{m, j}^c I_k(x)^m W_k^j(x) \right)$
 - $I^{k+1}(x) = \text{proj}_{[0, 1]} \left(\frac{I^k(x) + \Delta t \left(\alpha \mu + \beta I_0(x) + \frac{1}{2} \tilde{\mathcal{C}}_{w, k}^{[\varphi^c, n]}(x) \right)}{1 + \Delta t (\alpha + \beta)} \right)$
-

Note that it has been assumed that the contrast component $\sum_{j=0}^n \sum_{m=0}^{n-j} p_{m, j} I_k(x)^m (w * I_k^j)(x) \in [-1, 1]$. Due to numerical inaccuracies, it could be necessary to fix this function to ensure that

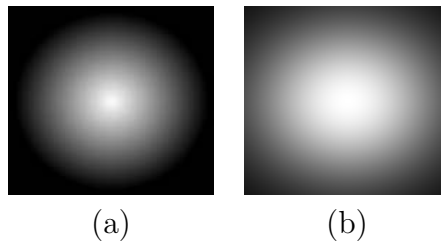


Figure 1: Two implemented kernels setting $\sigma = 3$ (a) linear (b) Gaussian.

$I_{k+1}(x)$ stays in the correct interval, which is done by projecting $I_{k+1}(x)$ into $[0, 1]$ and $\tilde{\mathcal{C}}_{w,k}^{[\varphi^c, n]}(x)$ into $[-1, 1]$ (function $\text{proj}_{[a,b]}(x)$ in Algorithm 2).

Weighting kernel w . The spatial kernel focuses the contrast increase of the contrast term around zero. The area of influence of the kernel is related to the final level of contrast enhancement. The more narrow the kernel, the more local the contrast increase. Here, we propose the use of two very different kernels:

- Euclidean:

$$w_\sigma(x, y) = C(1 - \sigma\sqrt{x^2 + y^2}) \tag{18}$$

- Gaussian:

$$w_\sigma(x, y) = C \exp\left\{-\frac{x^2 + y^2}{2\sigma^2}\right\} \tag{19}$$

where the constant C is set so the overall mass of the kernel adds to 1. Note that the effect of changing σ is opposite in both kernels. A bigger value of σ for the Euclidean kernel produces a narrower kernel, while incrementing it in the Gaussian kernel produces a wider kernel. In the implementation, to ease its use, we normalize this behaviour by inverting the σ value (the input value of σ is redefined as $1/\sigma$) for the *Gaussian* kernel. An example of both kernels can be observed in Figure 1.

Complexity of the final algorithm. The minimization process has now been reduced to a pixel-wise set of operations. Note that, for each k -iteration, there is no need to compute the convolution for each pixel x . On the contrary, the n convolutions $w * I_k^j$ can be *pre-computed* through the FFT (Fast Fourier Transform). Thus, the overall complexity of the algorithm drastically decreases to $\mathcal{O}(N \log N)$.

4 Results

In this section, we present some results obtained with the described algorithm, as well as the influence of the different parameters. We start with the parameters related to the contrast term, described in Section 3.2. Then, we detail the balancing terms α, β and finally, we describe a known problem of the algorithm, that is, the over-enhancement of noise when dealing with images acquired with low light intensity.

4.1 Parameters of the Contrast Term

Polynomial approximation. The first important parameter of the algorithm is the contrast function type, defined as parameter c in Algorithm 1. In Figure 2, we can see the three 2D-functions $\varphi^{\log}, \varphi^{-\mathcal{M}}$, and φ^{id} that the polynomial needs to approximate, as described in Section 3.2. The main

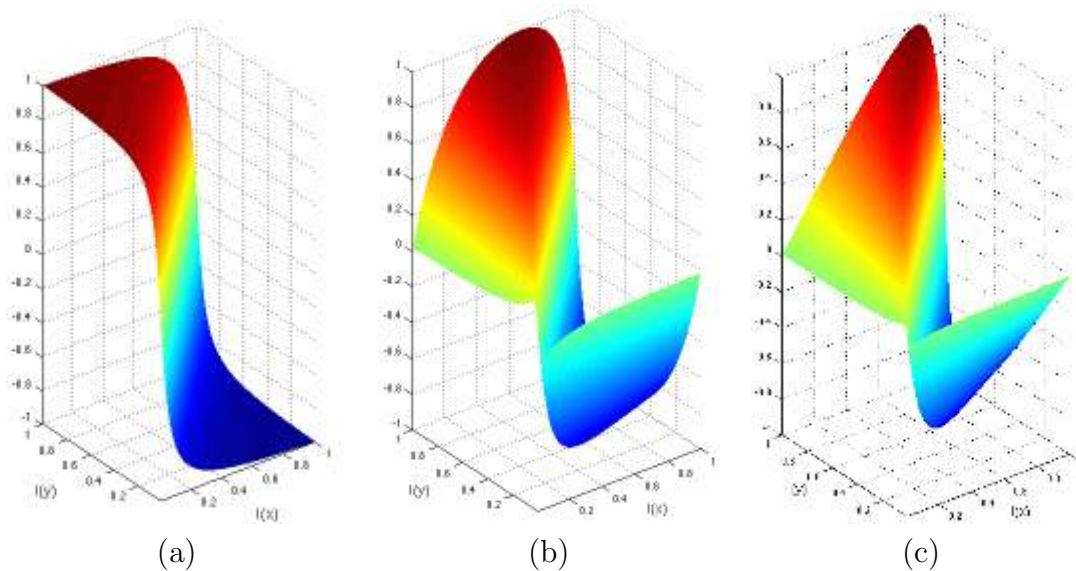


Figure 2: Surfaces obtained with Algorithm 1 for (a) $c = \log$ (b) $c = -\mathcal{M}$ (c) $c = \text{id}$.

difference between the functions relies in the extremes of the intensity range. In Figure 3 we show the impact on the same images of changing the polynomial.

Weighting function diameter σ . The impact of the contrast term $\tilde{\mathcal{C}}_{w,k}^{[\varphi^c,n]}$ depends on the area of influence of the kernel, which is defined by σ . As we increase σ , the contrast augments locally. Numerically, the values of σ are normalized to be adapted to the size of the input image. In Figure 4 we can see the effect, on the output images, of changing the kernel w and the parameter σ . The first row corresponds to the original images, and the outputs are shown in the following rows.

On the first image set, we would like to point out how the algorithm is reproducing a well known perceptual illusion known as “Mach bands” effect. In the original image, all bands are flat, but our visual system perceives them as having an “overshoot” to “undershoot” next to the intensity edges. This effect is reproduced by the output of the algorithm for $\sigma = 1$ for both kernels. This process implies that our perception is changing the intrinsic geometry of the image, since it is introducing new level lines. In fact, if we continue narrowing the kernel ($\sigma = 3$), the order of the level lines gets inverted, and the algorithm starts introducing an artifact called “sharpening”. This same behavior can be observed in a more subtle way on the results of the tree, specially on its leaves. The final set of images shows how the algorithm is able to discard a very strong color cast successfully with both types of kernels, due to the independent application of the algorithm to each color channel.

4.2 Global Parameters: α, β

The parameters α, β balance the influence of the attachment to the mean value and original input image, respectively. The bigger the values the more influence their term has in the final output, as can be observed in Figure 5. Note that for small values of α and β , the contrast term becomes more important.

4.3 Possible Artifacts

This algorithm increases locally the contrast of an image. Intrinsically, this process can introduce two types of artifacts:



Figure 3: Results obtained with the polynomials that approximate (a) $c = \log$ (b) $c = -\mathcal{M}$ (c) $c = \text{id}$. Note how the blue area of the book is rendered differently in the first set of images, as well as the contrast on the leaves and clouds on the tree images.

- Noise enhancement: When applied to pictures taken under very dim conditions or with strong *jpeg* blocking, these artifacts can become visible, as can be observed on the first row of Figure 6.
- Sharpening: As explained previously, it consists in inverting the order of the level lines as shown in Figure 5 or Figure 6. This effect should not necessarily be considered as an artifact, but rather as the result of the deliberate application of a non-realistic style.

5 Conclusions

In this article, we have presented a thorough explanation and test of the algorithm presented in [7]. This algorithm enhances contrast locally and improves the visibility of details in the input image. An interesting consequence of this process is that the global color cast produced by a colored light is removed from the image. The algorithm is able to successfully do these tasks because it implements several perceptual mechanisms developed by the human visual system: light adaptation, local contrast enhancement and color constancy. Although the method relies on a set of parameters, whose influence has been explored in this document, they can be set in a way that gives good results for most images.

Image Credits



Edoardo Provenzi, CC-BY-SA 3.0



Wiktionary (http://en.wiktionary.org/wiki/oak_tree) CC-BY-SA 3.0



Kobus Barnard, SFU Computational Vision Laboratory³



Source image courtesy of albatros11, licensed by Getty images⁴

³<http://www.cs.sfu.ca/~colour/data/>

⁴<https://www.flickr.com/photos/albatros11/2509508972>

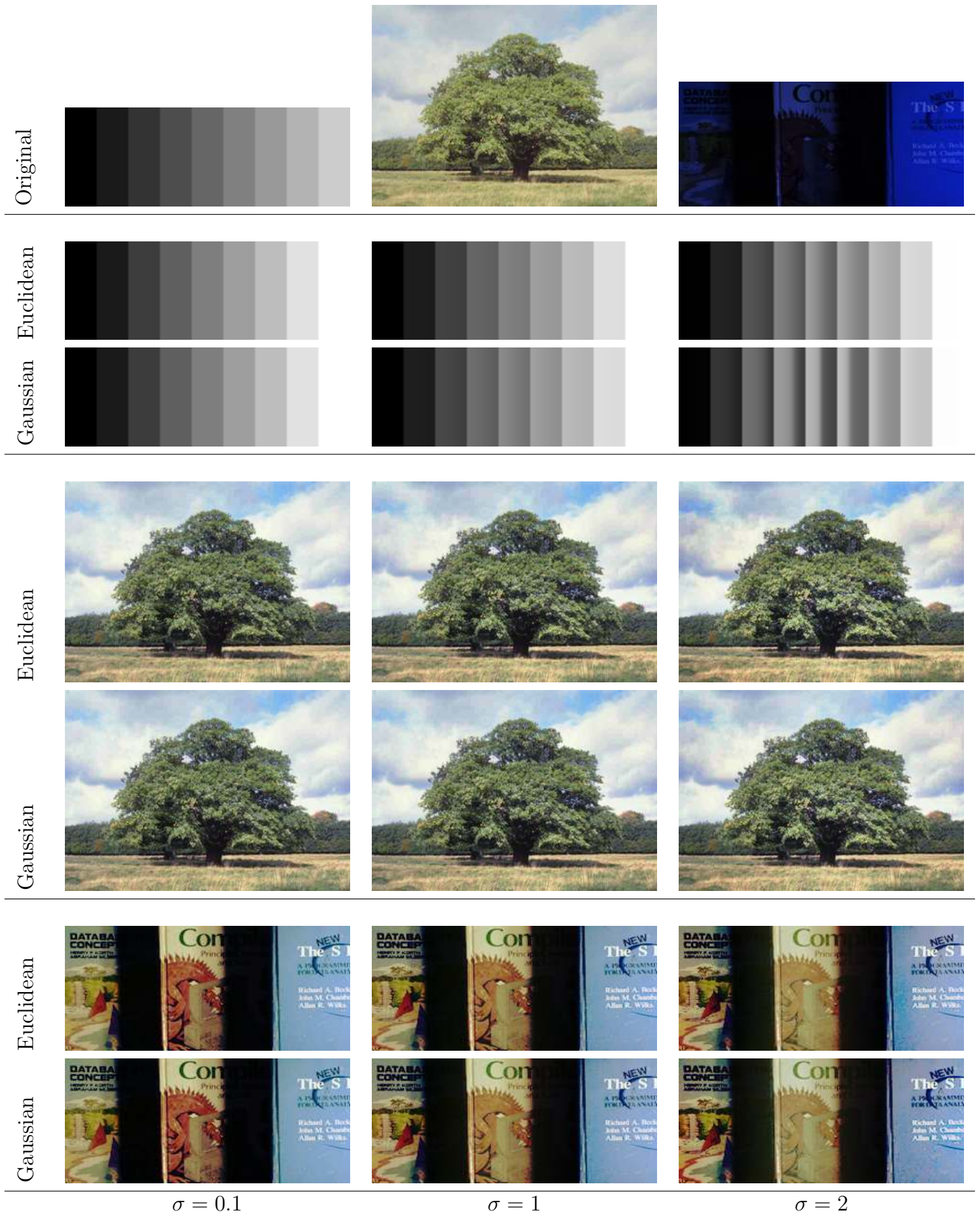


Figure 4: Results obtained by setting $c = \log$ in Algorithm 1, and $\alpha = 1.1$ and $\beta = 1$ in Algorithm 2. Note that the algorithm can produce an inversion of the level lines usually related to the sharpening effect, specially visible for $\sigma = 2$.



Figure 5: Changes in the output as α and β change. Note that for small values of α and β the strong term is the contrast term $\tilde{C}_{w,k}^{[\varphi^c, n]}(x)$ in Equation (12). Increasing α ensures the global condition that the mean is close to μ . Increasing β forces the result to be closer to the original image, which in this case is dark and with a blue color cast. For all these results, $c = \text{id}$, $\sigma = 1$ and w is Gaussian.

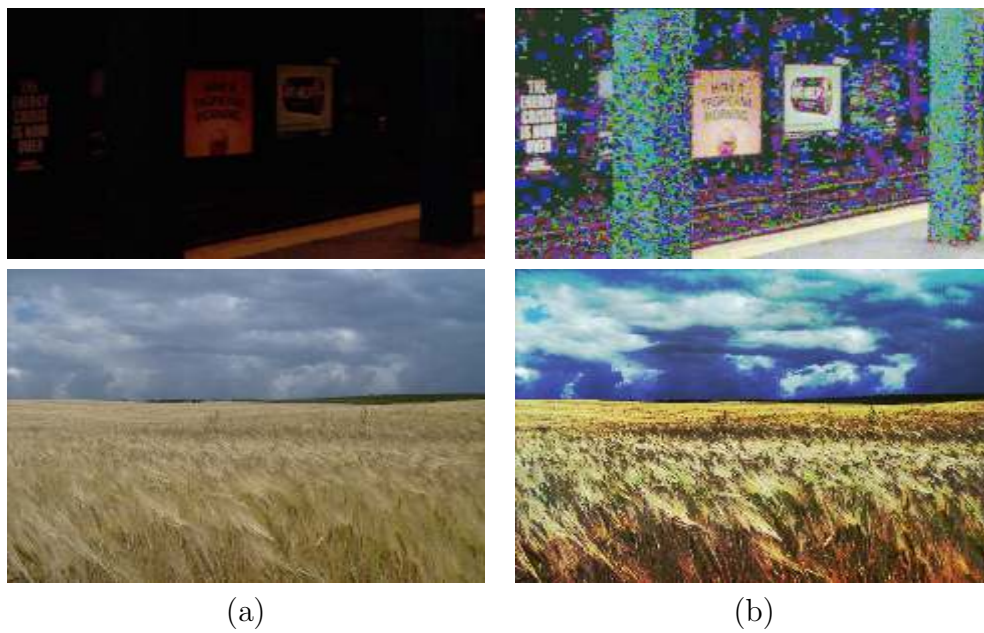


Figure 6: (a) Original image. (b) Result obtained with $c = \text{id}$, $\alpha = 1.1$, $\beta = 1$, w is Gaussian and $\sigma = 0.5$ for the first row and $\sigma = 3$ for the second row. This in an example where the noise and the jpeg blocks are made visible, or a sharpening is produced after using the described algorithm.

References

- [1] L.A. AMBROSIO, N. GIGLI, AND G. SAVARÉ, *Gradient flows in metric spaces and in the space of probability measures*, Lectures in Mathematics, Birkhauser, 2005.
- [2] M. BERTALMIÓ, V. CASELLES, AND E. PROVENZI, *Issues about the retinex theory and contrast enhancement*, International Journal of Computer Vision, (2009). <http://dx.doi.org/10.1007/s11263-009-0221-5>.
- [3] M. BERTALMIÓ, V. CASELLES, E. PROVENZI, AND A. RIZZI, *Perceptual color correction through variational techniques*, IEEE Transactions on Image Processing, 16 (2007), pp. 1058–1072. <http://dx.doi.org/10.1109/TIP.2007.891777>.
- [4] R.C. GONZALEZ AND R.E. WOODS, *Digital image processing*, Prentice Hall, 2002.
- [5] D.H. HUBEL, *Eye, Brain, and Vision*, Scientific American Library, 1995.
- [6] E.H. LAND AND J.J. MCCANN, *Lightness and Retinex theory*, Journal of the Optical Society of America, 61 (1971), pp. 1–11.
- [7] R. PALMA-AMESTOY, E. PROVENZI, M. BERTALMIÓ, AND V. CASELLES, *A perceptually inspired variational framework for color enhancement*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 31 (2009), pp. 458–474. <http://dx.doi.org/10.1109/TPAMI.2008.86>.
- [8] 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, 22 (2005), pp. 2613–2621. <http://dx.doi.org/10.1364/JOSAA.22.002613>.
- [9] 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*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 30 (2008), pp. 1757–1770. <http://dx.doi.org/10.1109/TPAMI.2007.70827>.
- [10] A. RIZZI, C. GATTA, AND D. MARINI, *A new algorithm for unsupervised global and local color correction*, Pattern Recognition Letters, 24 (2003), pp. 1663–1677. [http://dx.doi.org/10.1016/S0167-8655\(02\)00323-9](http://dx.doi.org/10.1016/S0167-8655(02)00323-9).
- [11] G. SAPIRO AND V. CASELLES, *Histogram modification via differential equations*, Journal of Differential Equations, 135 (1997), pp. 238–268. <http://dx.doi.org/10.1006/jdeq.1996.3237>.
- [12] R. SHAPLEY AND C. ENROTH-CUGELL, *Visual adaptation and retinal gain controls*, vol. 3, 1984, pp. 263–346.
- [13] G. WYSZECKY AND W. S. STILES, *Color science: Concepts and methods, quantitative data and formulas*, John Wiley & Sons, 2000.