Variational models for color image correction inspired by visual perception and neuroscience

Thomas Batard, Johannes Hertrich, Gabriele Steidl

▶ **To cite this version:**

Thomas Batard, Johannes Hertrich, Gabriele Steidl. Variational models for color image correction inspired by visual perception and neuroscience. 2020. hal-02463731v2

HAL Id: hal-02463731
https://hal.science/hal-02463731v2

Preprint submitted on 17 Jun 2020

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

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

Noname manuscript No.
(will be inserted by the editor)

Variational models for color image correction inspired by visual perception and neuroscience

Thomas Batard · Johannes Hertrich · Gabriele Steidl

Abstract Reproducing the perception of a real-world scene on a display device is a very challenging task which requires the understanding of the camera processing pipeline, the display process, and the way the human visual system processes the light it captures. Mathematical models based on psychophysical and physiological laws on color vision, named Retinex, provide efficient tools to handle degradations produced during the camera processing pipeline like the reduction of the contrast. In particular, Batard and Bertalmío [J Math. Imag. Vis. 60(6), 849-881 (2018)] described some psychophysical laws on brightness perception as covariant derivatives, included them into a variational model, and observed that the quality of the color image correction is correlated with the accuracy of the vision model it includes.

Based on this observation, we postulate that this model can be improved by including more accurate data on vision with a special attention on visual neuroscience here. Then, inspired by the presence of neurons responding to different visual attributes in the area V1 of the visual cortex as orientation, color or movement, to name a few, and horizontal connections modeling the interactions between those neurons, we construct two variational models to process both local (edges, textures) and global (contrast) features. This is an improvement with respect to the model of Batard and Bertalmío as the latter can not process local and global features independently and simultaneously. Finally, we conduct experiments on color images which corroborate the improvement provided by the new models.

Keywords Variational model · Differential geometry · Color image processing · Convex/nonconvex analysis · Visual neuroscience · Visual perception.

T. Batard (corresponding author)
Department Information and Communications Technologies
University Pompeu Fabra, 08003 Barcelona, Spain
E-mail: thomas.batard@upf.edu

J. Hertich
Department of Mathematics, TU Berlin, 10587 Berlin, Germany
E-mail: j.hertrich@math.tu-berlin.de

G. Steidl
Department of Mathematics, TU Berlin, 10587 Berlin, Germany
E-mail: steidl@math.tu-berlin.de

1 Introduction

1.1 On existing models aiming to reproduce the perception of real-world scenes on standard monitors and their limits

Due to physical and technological limitations of the acquisition process of a real-world scene by a digital camera and its reproduction on a display device, the image perceived on a monitor appears as a degraded version of the original scene. The degradations observed are, for instance, the presence of noise, mismatch of colors, and loss of details and contrast to name only a few [11].

Image enhancement can be defined as the procedure of improving the quality and information content of an image for display purpose or for further tasks like analysis and detection. Image enhancement techniques, like denoising, deblurring, contrast stretching, can then be used to undo the degradations generated during the image acquisition process. Among image enhancement techniques, we distinguish in this paper the ones which increase the visibility of image features, which we call enhancement techniques too, from the ones which decrease the visibility of image features, which we call regularization techniques.

Nonetheless, as the dynamic range of a real-world

scene is much higher than the one of a standard display device, processing the camera output image so that it reproduces the appearance of the original real-world scene is even more challenging than recovering its light intensity. One way to tackle the problem is to make use of knowledge about the functioning of the human visual system (HVS). The seminal model based on this approach, named Retinex, aimed to reproduce the perception of scenes using psychophysical laws on color perception [27],[28]. However, it suffers from drawbacks when applied to digital images, like the systematic increase of the intensity values of the image as proved in [32] and the generation of noise and artifacts. Since then, many models inspired by Retinex have been developed with the aim of overcoming the drawbacks of the original formulation (see e.g. [10],[41] and references therein). These models are commonly also called Retinex models.

A connection between Retinex models and visual neuroscience has been established for the first time by Cowan and Bressloff [16] through the Wilson-Cowan (W-C) equations [39],[40], a system of ODEs describing the temporal evolution of the activity of a population of neurons responding to some visual attributes and located in the area V1 of the visual cortex. Later on, Bertalmío et al. [8] proposed a variational model for color image correction which performs color cast removal and contrast enhancement and improves the results of the original Retinex formulation. They showed that this model is related to the W-C equations by showing that the gradient descent of a differentiable approximation of the model can be identified with a simplified version of the W-C equations. Then, Bertalmío and Cowan [9] connected the variational model for color image correction [8], the so-called kernel-based Retinex [10], and the simplified W-C equations, establishing a connection between image processing, visual perception and visual neuroscience.

More recently, Batard and Bertalmío [6] proposed a variational model on vector bundles for color image correction which extends the model of Bertalmío et al. [8] in the sense that the latter can be viewed as the Euclidean restriction of the former. In this context, a color image is considered as a section of a vector bundle of rank 3 over the image domain, and the image gradient is encoded into a covariant derivative. This approach is consistent with Georgiev's approach of color perception [21] in which he suggests that the color constancy property of the HVS is due to the equivariance property of the perceived gradient of the observed scene with respect to lighting changes, and which turns the perceived gradient of an image into a covariant derivative. Moreover, it has been shown in [6] that some psychophysical laws like Weber's and Weber-Fechner's laws, and the Helmholtz-Kohlrausch effect arise from connection 1-forms which are solutions of some variational models constructed in [4]. This suggests that the HVS processes the light in an optimal manner.

Despite the fact that the image processing models of Bertalmío et al. [8] and Batard and Bertalmío [6] are connected to visual perception and visual neuroscience, they still suffer from many limitations when trying to reproduce the perception of real-world scenes. Examples of these limitations are:

1. The models only deal with very few perceptual attributes, i.e. the contrast (global feature) and edges, textures, noise (local features) and they are not able to process global and local features both independently and simultaneously. Whereas the parameters chosen in the model of Bertalmío et al. make the model only process the contrast, Batard and Bertalmío use a parameter which controls the locality of the features that are processed, where a large value implies a processing of the contrast, a small one a processing of the local features, and a medium value provides a trade-off between local and global features.

2. The models systematically enhance the image features. Whereas enhancing the image features (e.g. edges, textures, contrast) can be desirable when the image to be processed is of low-resolution, e.g. it is the output image of a low-resolution camera, recent progresses in technology and imaging, e.g. High Dynamic Range (HDR) imaging [33], can provide images of great quality (high contrast, saturated colors, etc). It can even occur that these images, when displayed on standard monitors, have higher contrast than the original real-world scenes, for instance if the tone mapping method which enables the HDR image to be displayed on a standard monitor is not well adjusted [17],[19]. In such a case, an accurate color image correction model should be able to regularize the image features in order to match the perception of the real-world scene.

1.2 Using neuroscience to overcome drawbacks of color image correction models

The two limits of the variational models aforementioned can be overcome by making use of some properties of the functional architecture of the HVS related to the W-C equations.

1. The W-C equations can describe the temporal activity of neurons responding to various visual attributes.

Whereas the original W-C equations deal with neurons responding to local orientations, it has been suggested in [15] resp. [36] that the activities of neurons responding to local features (edges and textures) resp. colors evolve according to W-C equations as well.
Indeed, in the area V1 of the visual cortex, the presence of neurons responding to visual attributes, like color, orientation, movement, etc, and arranged in a hypercolumnar structure has been discovered by Hubel and Wiesel [23],[24]. As a consequence, the set of neurons responding to such a visual attribute can be represented as a fiber bundle $\Omega \times \mathcal{F} \longrightarrow \Omega$, where Ω denotes the cortical plane and $\mathcal{F}$ a set representing the visual attribute (e.g. $\mathcal{F}$ is an opponent space if the visual attribute considered are the colors). Then, according to the W-C equations, the activities $a(x, f, t)$ and $a(y, f', t)$ at the time t of the neurons of coordinates (x, f) and (y, f') in $\Omega \times \mathcal{F}$ are connected through a term of the form

$$w(x, f, y, f')\, \sigma\left[s\left(a(x, f, t), a(y, f', t)\right)\right], \tag{1}$$

where $w(x, f, y, f')$ is called the synaptic weight between the neurons (x, f) and (y, f'), σ is an activation function such that $\sigma(0) = 0$, and s measures a difference.
When establishing a connection between their image processing model and the W-C equations, Bertalmío et al. [8] deal with a simplified version of the term (1) by removing its dependency with respect to the visual attribute, yielding a term of the form

$$w(x, y)\sigma\left[s\left(a(x, t), a(y, t)\right)\right] \tag{2}$$

under the identification between the cortical plane and the image domain. Then, in order for the term (2) to characterize a visual attribute, the couple of parameters (w, σ) has to be carefully chosen.
In [5], it has been shown that, there exists a choice (w_1, σ_1) which makes the term (2) characterize edges and textures (local features). On the other hand, it has been shown in [7] that there exists a choice (w_2, σ_2) which makes the term (2) characterize the contrast (global feature), such that $w_2 \neq w_1$ and $\sigma_2 \neq \sigma_1$. Based on this, we propose to combine these two terms in order to process independently and simultaneously the local and global features. Note that it has been shown in [6] that a well-chosen difference measure s can describe some psychophysical laws.

2. Horizontal connections between columns can have either excitatory or inhibitory effect. In the W-C equations, the synaptic weight (1) is decomposed as

$$\begin{aligned} w(x, f, y, f') = {} & w_{ver}(f, f')\delta(x - y) \\ & + \gamma\, w_{hor}(x, f, y, f')\,(1 - \delta(x - y)), \end{aligned}$$

where δ is the Dirac delta function, $\gamma \in \mathbb{R}$, w_{ver} is called vertical connection and w_{hor} horizontal connection, see [39,40,15,36] for more details about the shape of the vertical and horizontal connections. The sign of γ determines the effect of the horizontal connection, excitatory or inhibitory, and which actually depends on the contrast in the input signal [13]. Under the simplification aforementioned, the term (2) takes the form

$$\gamma\, w_{hor}(x, y)\sigma\left[s\left(a(x, t), a(y, t)\right)\right], \tag{3}$$

and the sign of γ determines whether the term (3) has a regularizing or enhancing effect in the subsequent image processing models. Hence, attaching a parameter γ_1 to the term describing the local features and a parameter γ_2 to the term describing the global features makes the proposed models regularize or enhance the local features and regularize or enhance the global features depending on the signs of γ_1 and γ_2.

1.3 Contribution of the paper: construction and analysis of two models processing the local and global features independently and simultaneously

Because the way the visual input is processed in V1 is very complex and still not fully understood, we present in this paper two models for color image correction based on two different hypothesis. The first model relies on the idea that chromatic and achromatic information interact as early as V1 [26], leading to a variational model composed of three terms: a data term, a term characterizing edges and textures, and a term characterizing the contrast. The second model is based on the idea that achromatic and chromatic components are processed separately and independently in early vision and only interact at some high perceptual level beyond V1, as it has been originally proposed by Hering [25] in the 19th century. Together with the idea that neurons responding to local features are activated by the achromatic information and neurons responding to colors are activated by the chromatic information of the visual input, we are led to consider as a second model a system of two variational models: one similar to the model in [7] to process the contrast of the chrominance component, and one similar to the model in [6] to process the local features of the achromatic component of a color image.

The outline of the paper is the following. In section 2, we establish a connection between the image processing model [6] and the W-C equations for neurons responding to edges and textures [15], and between the model in [7] and the W-C equations for neurons responding to colors [36]. In section 3, we introduce the

new models, from which we derive new forms of the W-C equations. Then, we study their properties and propose a numerical scheme to reach their solutions. Finally, experiments are conducted in section 4, where we show the capability of the proposed models to correct different degradations on color images.

2 Connection of W-C equations to vision and image processing

2.1 Mathematical setting

In order to analyze the variational models we introduce later on, we first need to remind some mathematical concepts.

Let E be a vector bundle over a manifold M with tangent bundle TM and cotangent bundle T^*M. By $\Gamma(E)$ we denote the set of continuous sections of E and by $\mathrm{pr}_1(E)$ the vector bundle over $M \times M$ induced by the projection

$$\mathrm{pr}_1 : M \times M \longrightarrow M, \qquad (x, y) \longmapsto x,$$

i.e.

$$\mathrm{pr}_1(E) = \{(x, y, p) \in M \times M \times E : x = \pi(p)\},$$

where π stands for the projection from the fibers to the base manifold.

Let ∇ be a covariant derivative on E and ω the corresponding connection 1-form, i.e. $\nabla = d + \omega$, where d stands for the standard differential operator and $\omega \in \Gamma(T^*M \otimes \mathrm{End}(E))$. By τ we denote the parallel transport map induced by ∇.

A *nonlocal covariant derivative* induced by ∇ is an operator $\nabla^{NL}_w \colon \Gamma(E) \longrightarrow \Gamma(\mathrm{pr}_1(E))$ defined by

$$\nabla^{NL}_w a \colon (x, y) \longmapsto w(x, y) \left(\tau_{y,x,\gamma_{y,x}} a(y) - a(x)\right),$$

where w is a smooth, positive, symmetric and normalized kernel as in (8), and $\tau_{y,x,\gamma_{y,x}}$ stands for the parallel transport along the path $\gamma_{y,x}$ joining y and x.

A positive definite metric h on E induces an L^2 scalar product and an L^p norm on $\Gamma(E)$, given by

$$\langle a_1, a_2 \rangle_{L^2(E)} = \int_M (a_1(x), a_2(x))_{h(x)} \, dx,$$

and

$$\|a\|_{L^p(E)} = \left(\int_M \|a(x)\|^p_{h(x)} \, dx \right)^{1/p}.$$

The space $L^p(E)$ is defined as the completion of $\Gamma(E)$ with respect to $\| \, . \, \|_{L^p(E)}$.

Moreover, h induces an L^2 scalar product and a L^p norm on $\Gamma(\mathrm{pr}_1(E))$ given by

$$\langle \zeta_1, \zeta_2 \rangle_{L^2(\mathrm{pr}_1(E))} = \int_{M^2} (\zeta_1(x, y), \zeta_2(x, y))_{h(x)} \, dx \, dy,$$

and

$$\|\zeta\|_{L^p(\mathrm{pr}_1(E))} = \left(\int_{M^2} \|\zeta(x, y)\|^p_{h(x)} \, dx \, dy \right)^{1/p}.$$

The space $L^p(\mathrm{pr}_1(E))$ is defined as the completion of $\Gamma(\mathrm{pr}_1(E))$ with respect to $\| \, . \, \|_{L^p(\mathrm{pr}_1(E))}$.

The adjoint of the operator ∇^{NL}_w is defined as the operator $\nabla^{NL^*}_w \colon \Gamma(\mathrm{pr}_1(E)) \longrightarrow \Gamma(E)$ satisfying

$$\langle \nabla^{NL^*}_w \zeta, \chi \rangle_{L^2(E)} = \langle \zeta, \nabla^{NL}_w \chi \rangle_{L^2(\mathrm{pr}_1(E))}$$

for all $\zeta \in \Gamma(\mathrm{pr}_1(E)), \chi \in \Gamma(E)$. In particular, if ∇ is compatible with h, i.e. if it satisfies

$$d_v h(a_1, a_2) = \langle \nabla_v a_1, a_2 \rangle_{L^2(E)} + \langle a_1, \nabla_v a_2 \rangle_{L^2(E)},$$

for all $v \in \Gamma(TM)$, then we have

$$\nabla^{NL^*}_w \zeta \colon x \longmapsto \int_M w(x, y) \left[\tau_{y,x,\gamma_{y,x}} \zeta(y, x) - \zeta(x, y)\right] \, dy,$$

see [5] for details of the computation.

Finally, by [5], it holds

$$\|\nabla^{NL}_w a\|_{L^1(E)} = \sup_{\zeta \in K_1} \langle a, \zeta \rangle_{L^2(E)}, \tag{4}$$

where $K_1 \colon =$

$$\overline{\left\{\nabla^{NL^*}_w \eta \colon \eta \in \Gamma(\mathrm{pr}_1(E)), \|\eta(x, y)\|_{h(x)} \leq 1 \ \ \forall x, y \in M\right\}}.$$

From now on, we consider the image domain, which we identify to the cortical plane in this paper, as a compact subset Ω of $\mathbb{R}^2$, and we denote by E the vector bundle $\Omega \times \mathbb{R}^3 \longrightarrow \Omega$. Moreover, we assume that E is equipped with a positive definite metric h and a covariant derivative compatible with h. A color image is considered as an element of the space $\Gamma(E) \cap L^2(E)$. For shortness, we will write $L^2(E)$ instead of $\Gamma(E) \cap L^2(E)$.

2.2 W-C equations and visual perception

The W-C equations describe the temporal evolution of the mean activity $a(x,f,t)$ of a population of neurons of cortical coordinates $x \in \Omega$ and feature preference $f \in \mathcal{F}$ with the following integro-differential equation

$$\frac{\partial a(x,f,t)}{\partial t} = -\alpha_1\, a(x,f,t) + a_0(x,f,t) + \alpha_2 \int_\Omega \int_{\mathcal{F}} w(x,f,y,f')\, \sigma\left[s\left(a(x,f,t), a(y,f',t)\right)\right]\, dy\, df', \tag{5}$$

where $\alpha_1, \alpha_2 > 0$ and a_0 is the visual input. The other terms have already been described in sect.1.2. In the case where the visual attribute corresponds to the local features (edges and textures), Chossat and Faugeras [15] proposed to encode the features space $\mathcal{F}$ with the set of structure tensors $\mathcal{T} \in \mathrm{SPD}(2,\mathbb{R})$, see [20]. Here, $\mathrm{SPD}(2,\mathbb{R})$ denotes the cone of symmetric positive definite matrices equipped with the affine invariant metric and corresponding distance function $d_{\mathrm{SPD}(2,\mathbb{R})}$. In the case where the visual attribute is the colors, Song et al. [36] proposed to encode $\mathcal{F}$ with some color opponent space $\mathfrak{C}_{opp} \subset \mathbb{R}^3$.

Ignoring the dependence of the neuron activity with respect to the feature, the term (1) is reduced to (3), and assuming that the input a_0 is constant in time, the W-C equation (5) is reduced to

$$\frac{\partial a(x,t)}{\partial t} = -\alpha_1\, a(x,t) + a_0(x) + \alpha_2 \gamma \int_\Omega w_{hor}(x,y) \sigma\left[s\left(a(x,t), a(y,t)\right)\right]\, dy. \tag{6}$$

It turns out that there is a strong analogy between the integral term in (6) and the Retinex theory of visual perception (see Appendix A.1). Moreover, the perceptual difference (with respect to some psychophysical law) between the pixels x and y can be written as

$$a(x) - \tau_{y,x,\gamma_{y,x}} a(y),$$

for a well-chosen metric h and a covariant derivative compatible with h (see Appendices A.2 and A.3). As a consequence, the notion of perceptual difference can be inserted into the simplified W-C equation (6) under the identification

$$s(a(x,t), a(y,t)) = a(x,t) - \tau_{y,x,\gamma_{y,x}} a(y,t). \tag{7}$$

2.3 W-C equations and image processing

From now on, we assume that $\alpha_1 = \alpha_2 = 1$ in (6), that the difference function s is the perceptual distance (7), and that w_{hor} is smooth, positive, symmetric, and normalized as

$$\int_\Omega w_{hor}(x,y)\, dy = \int_\Omega w_{hor}(x,y)\, dx = 1 \tag{8}$$

for all $x, y \in \Omega$.

In this paper, we assume that color images are expressed in the CIE $L^*a^*b^*$ color space. This choice is motivated by the connections between the proposed image processing models and color science established in this paper. Indeed, in sect. 2.3.2, we establish a connection with the W-C equations for neurons responding to colors in which an opponent color space is considered (see sect. 2.2). Moreover, the covariant derivative described in Appendix A.3., which we consider in several experiments, is closely related to the brightness defined in [18] and expressed in the CIE $L^*a^*b^*$ space.

Note that, in [8], the authors relate their image processing model to the original W-C equations for neurons responding to local orientations (see sect. 2.2). However, this model processes the contrast of the image and not its local orientations. Hence, we claim that it would be more accurate to relate this image processing model to the W-C equations for neurons responding to colors.

2.3.1 Image processing model related to the W-C equations for neurons responding to edges and textures

Assuming that the activation function σ in (6) is nonlinear of the form

$$\sigma(z) = \frac{z}{\sqrt{\|z\|_h^2 + \epsilon}}, \quad \epsilon > 0,$$

the evolution equation (6), together with the perceptual difference (7), corresponds to the gradient descent equation of a differentiable approximation of the variational problem

$$\underset{a \in L^2(E)}{\arg\min}\ \frac{1}{2}\int_\Omega \|a(x) - a_0(x)\|_{h(x)}^2\, dx - \frac{\gamma}{2}\int_{\Omega^2} w_{hor}(x,y) \| \tau_{y,x,\gamma_{y,x}} a(y) - a(x)\|_{h(x)}\, dx\, dy. \tag{9}$$

The term

$$\| \tau_{y,x,\gamma_{y,x}} a(y) - a(x)\|_{h(x)}$$

can be considered as the perceptual distance between the pixels x and y. Note that by compatibility of the covariant derivative with h, we have

$$\| \tau_{y,x,\gamma_{y,x}} a(y) - a(x)\|_{h(x)} = \| \tau_{x,y,\gamma_{x,y}} a(x) - a(y)\|_{h(y)}.$$

Model (9) has been first studied in [5], where the uniqueness of the solution has been proved for $\gamma \leq 0$ and the existence of solutions for $\gamma > 0$ and Ω discrete. Moreover, it has been shown that the solutions $\underline{a}$ can be written as

$$\underline{a} = a_0 - \begin{cases} \underset{a^* \in K_{-\gamma/2}}{\arg\min} \left\| a_0 - a^* \right\|^2_{L^2(E)} & \text{if} \quad \gamma \leq 0, \\ \underset{a^* \in K_{\gamma/2}}{\arg\max} \left\| a_0 - a^* \right\|^2_{L^2(E)} & \text{if} \quad \gamma > 0 \\ & \text{and} \quad \Omega \quad \text{is discrete.} \end{cases}$$

In particular, taking w_{hor} as the kernel (here discrete)

$$w_{hor}(x,y) = \begin{cases} \dfrac{1}{|\Omega'(x)|} & \text{if} \quad y \in \Omega'(x), \\ 0 & \text{otherwise,} \end{cases} \tag{10}$$

where $\Omega'(x)$ is a medium-sized window centered at x, the model (9) processes the local features of a_0. For an example, see Fig. 1(center), where the model has been applied with h given by the identity matrix and $\omega = 0$ in the frame $L^*a^*b^*$. As parameters $\gamma = 10$, and Ω' of size 41×41 were chosen.
Hence, by its property of processing the local features of images, together with the observation described in the paragraph following formula (8), model (9) with w_{hor} of the form (10) can be related to the W-C equations describing the activity of neurons responding to the local features of the visual input, i.e. equation (5) with $\mathcal{F}$=SPD(2,$\mathbb{R}$).

2.3.2 Image processing model related to the W-C equations for neurons responding to colors

Assuming that the activation function σ is the Identity map, the evolution equation (6), together with the perceptual difference (7), corresponds to the gradient descent equation of the variational model

$$\underset{a \in L^2(E)}{\arg\min} \ \frac{1}{2} \int_\Omega \|a(x) - a_0(x)\|^2_{h(x)} \, dx \tag{11}$$
$$- \frac{\gamma}{4} \int_{\Omega^2} w_{hor}(x,y) \| \tau_{y,x,\gamma_{y,x}} a(y) - a(x) \|^2_{h(x)} \, dx \, dy,$$

which has been introduced in [7], but not studied, and we propose in what follows to analyze it. We obtain the following result.

Proposition 1 *Let $\Omega = \mathbb{T}^2$, where $\mathbb{T}$ denotes the 1-torus, i.e., we assume periodic boundary conditions. The energy in* (11) *is strictly convex for $\gamma < 1/2$. Moreover, for $w_{hor}(x,y) = w_{hor}(x-y)$, if the covariant derivative is flat and if we denote by P a moving frame in which the corresponding connection 1-form vanishes, then the unique solution of the problem* (11) *is given by*

$$P\left(\mathcal{F}^{-1} \left(\frac{\mathcal{F}(P^{-1} a_0)}{(1-\gamma) + \gamma \, \mathcal{F}(w_{hor})} \right) \right), \tag{12}$$

where $\mathcal{F}$ is the Fourier transform on $L^2(\mathbb{T}^2)$, and $\mathcal{F}^{-1}$ its inverse.

Proof Denoting by X the functional in (11), we have

$$\frac{d^2}{dt^2} X(a + t\psi)_{|t=0} = \int_\Omega \|\psi(x)\|^2_{h(x)} \, dx \tag{13}$$
$$- \frac{\gamma}{2} \int_{\Omega^2} w_{hor}(x,y) \|\tau_{y,x,\gamma_{y,x}} \psi(y) - \psi(x)\|^2_{h(x)} dx \, dy.$$

If $\gamma \leq 0$, then we observe that the expression (13) is positive $\forall \psi$ and vanishes if and only if $\psi = 0$.
Let us assume that $\gamma > 0$. From the triangle inequality and the compatibility of the covariant derivative with h, we have

$$\|\tau_{y,x,\gamma_{y,x}} \psi(y) - \psi(x)\|^2_{h(x)} \leq \|\psi(x)\|^2_{h(x)} + \|\psi(y)\|^2_{h(y)} + 2\|\psi(x)\|_{h(x)} \|\psi(y)\|_{h(y)}.$$

Moreover, using Fubini's theorem and the fact that w_{hor} is normalized, we have

$$\begin{aligned} &\int_{\Omega^2} w_{hor}(x,y) \|\psi(x)\|^2_{h(x)} dx dy \\ &= \int_{\Omega^2} w_{hor}(x,y) \|\psi(y)\|^2_{h(y)} dx dy \\ &= \int_\Omega \|\psi(x)\|^2_{h(x)} dx. \end{aligned}$$

Hence,

$$\begin{aligned} &\int_{\Omega^2} w_{hor}(x,y) \|\tau_{y,x,\gamma_{y,x}} \psi(y) - \psi(x)\|^2_{h(x)} dx \, dy \\ &\leq 2 \int_\Omega \|\psi(x)\|^2_{h(x)} dx \\ &\quad + 2 \int_{\Omega^2} w_{hor}(x,y) \|\psi(x)\|_{h(x)} \|\psi(y)\|_{h(y)} \, dx dy. \end{aligned} \tag{14}$$

Applying the Cauchy-Schwarz inequality yields

$$\begin{aligned} &\int_{\Omega^2} w_{hor}(x,y) \|\psi(x)\|_{h(x)} \|\psi(y)\|_{h(y)} \, dx dy \\ &\leq \left(\int_{\Omega^2} w_{hor}(x,y) \|\psi(x)\|^2_{h(x)} dx dy \right)^{1/2} \\ &\quad \times \left(\int_{\Omega^2} w_{hor}(x,y) \|\psi(y)\|^2_{h(y)} dx dy \right)^{1/2} \\ &\leq \int_{\Omega^2} w_{hor}(x,y) \|\psi(x)\|^2_{h(x)} dx dy \\ &\leq \int_\Omega \|\psi(x)\|^2_{h(x)} dx. \end{aligned}$$

Fig. 1 Different activations and kernels imply the processing of different features. Left: Original image - Middle: Solution of the model (9) (non-linear activation) with the kernel w_{hor} taken as a constant kernel of medium-size support - Right: Solution of the model (11) (linear activation) with the kernel w_{hor} taken as a Gaussian kernel of large variance.

Together with (14), it gives

$$\int_{\Omega^2} w_{hor}(x,y)\|\tau_{y,x,\gamma_{y,x}}\psi(y)-\psi(x)\|^2_{h(x)}dx\,dy \qquad (15)$$
$$\leq 4\int_{\Omega}\|\psi(x)\|^2_{h(x)}\,dx,$$

from which we deduce that

$$\frac{d^2}{dt^2}X(a+t\psi)_{|t=0}\geq(1-2\gamma)\int_{\Omega}\|\psi(x)\|^2_{h(x)}\,dx,$$

which means that X is strictly convex if $\gamma<1/2$.

The critical points of X are the solutions of the following gradient equation in the L^2 sense

$$a-a_0-\gamma\left(a-\int_{\Omega}w_{hor}(x,y)\tau_{y,x,\gamma_{y,x}}a(y)\,dy\right)=0. \quad (16)$$

Assume that $w_{hor}(x,y)=w_{hor}(x-y)$. In particular, if the covariant derivative is flat, and denoting by P the moving frame in which the corresponding connection 1-form vanishes, equation (16) is expressed in the frame P as

$$\mathbf{a}-P^{-1}a_0-\gamma(\mathbf{a}-w_{hor}*\mathbf{a})=0,$$

where $*$ denotes the cyclic convolution and $a(x)=P(x)\mathbf{a}(x)$. The unique solution of the above equation is given by

$$\mathbf{a}=\mathcal{F}^{-1}\left(\frac{\mathcal{F}(P^{-1}a_0)}{(1-\gamma)+\gamma\,\mathcal{F}(w_{hor})}\right).$$

Finally, the solution is expressed in the original frame by (12). □

Taking w_{hor} as a (periodized) Gaussian kernel of large variance, model (11) processes the contrast of a_0 (see Fig. 1(right), where the model has been applied with h given by the identity matrix $\mathbb{I}_3$ and $\omega=0$ in the frame $L^*a^*b^*$, and with the parameters $\gamma=0.25$ and variance of the Gaussian 2000). In particular, we observe that only the colors of the original image are processed and not its local features. Hence, the model (11) with w_{hor} being a Gaussian kernel of large variance can be related to the W-C equation (5) with $\mathcal{F}$ being an opponent color space $\mathfrak{C}_{opp}$, as the model in [8].

3 Variational models to process local and global features simultaneously and independently

3.1 Different hypothesis on the processing of the visual input in V1 yield different variational problems

Based on the results shown in Fig. 1, we propose to combine the models (9) and (11) in order to process local and global features simultaneously and independently.

3.1.1 Achromatic and chromatic components interact in V1

Under this hypothesis, we propose the following variational model

$$\begin{aligned}\arg\min_{a\in L^2(E)}\ &\frac{1}{2}\int_{\Omega}\|a(x)-a_0(x)\|^2_{h(x)}\,dx\\ &-\frac{\gamma_1}{2}\int_{\Omega^2}w_1(x,y)\,\left\|\tau_{y,x,\gamma_{y,x}}a(y)-a(x)\right\|_{h(x)}\,dx\,dy\\ &-\frac{\gamma_2}{4}\int_{\Omega^2}w_2(x,y)\,\left\|\tau_{y,x,\gamma_{y,x}}\,a(y)-a(x)\right\|^2_{h(x)}\,dx\,dy,\end{aligned} \qquad (17)$$

where w_1 is of the form (10), w_2 is a normalized Gaussian kernel, and $\gamma_1,\gamma_2\in\mathbb{R}$.

Following the connection between the W-C equations and image processing models established in sect. 2, model (17) can be associated to a new form of the W-C equations combining the activity of neurons responding to edges/textures and colors, which we write $a(x,\mathcal{T},c,t)$, for $x\in\Omega,\mathcal{T}\in\mathrm{SPD}(2,\mathbb{R})$, $c\in\mathfrak{C}_{opp}$, and $t\in\mathbb{R}^+$. It gives the following integro-differential equa-

tion

$$\frac{\partial a(x,\mathcal{T},c,t)}{\partial t} = -\alpha_1\, a(x,\mathcal{T},c,t) + a_0(x,\mathcal{T},c,t) + \alpha_2 \int_\Omega \left(\int_{\mathfrak{C}_{opp}} w_2(x,c,y,c') \sigma_2[s\,(a(x,\mathcal{T},c,t), a(y,\mathcal{T}',c',t))] dc' + \int_{\mathrm{SPD}(2,\mathbb{R})} w_1(x,\mathcal{T},y,\mathcal{T}') \sigma_1[s\,(a(x,\mathcal{T},c,t), a(y,\mathcal{T}',c',t))]\, d\mathcal{T}' \right) dy,$$

where the synaptic weights w_1, w_2 are of the form

$$w_1(x,\mathcal{T},y,\mathcal{T}') = w_{1_{ver}}(\mathcal{T},\mathcal{T}')\delta(x-y) + \gamma_1\, w_{1_{hor}}(x,\mathcal{T},y,\mathcal{T}')\,(1-\delta(x-y)), \tag{18}$$

and

$$w_2(x,c,y,c') = w_{2_{ver}}(c,c')\delta(x-y) + \gamma_2\, w_{2_{hor}}(x,c,y,c')\,(1-\delta(x-y)). \tag{19}$$

3.1.2 Achromatic and chromatic components are processed separately and independently in V1

Under this hypothesis, and together with the idea that neurons responding to local features are related to the achromatic component L^* and neurons responding to colors to the chrominance component $C^* = (a^*, b^*)$, we decompose the vector bundle $E = \Omega \times \mathbb{R}^3 \longrightarrow \Omega$ as $E = E_1 \oplus E_2$ with E_1 is the vector bundle $\Omega \times \mathbb{R} \longrightarrow \Omega$ equipped with a definite positive metric h_1 and a covariant derivative ∇_1 compatible with h_1, and E_2 is the vector bundle $\Omega \times \mathbb{R}^2 \longrightarrow \Omega$ equipped with a definite positive metric h_2 and a covariant derivative ∇_2 compatible with h_2. Then, denoting by L_0^* resp. C_0^* the achromatic resp. chromatic component of the image a_0, we propose the following system of variational model

$$\begin{cases} \underset{L^*\in L^2(\Omega;E_1)}{\arg\min}\ \dfrac{1}{2}\|L^*(x) - L_0^*(x)\|^2_{L^2(E_1)} - \dfrac{\gamma_1}{2}\|\nabla^{NL}_{1\,w_1} L^*\|_{L^1(E_1)}, \\ \underset{C^*\in L^2(\Omega;E_2)}{\arg\min}\ \dfrac{1}{2}\|C^*(x) - C_0^*(x)\|^2_{L^2(E_2)} - \dfrac{\gamma_2}{4}\|\nabla^{NL}_{2\,w_2} C^*\|^2_{L^2(E_2)}, \end{cases} \tag{20}$$

where w_1 is of the form (10), w_2 is a normalized Gaussian kernel, and $\gamma_1, \gamma_2 \in \mathbb{R}$.

The system (20) can be associated to the following system of W-C equations

$$\begin{cases} \dfrac{\partial a_1(x,\mathcal{T},t)}{\partial t} = -\alpha_{11} a_1(x,\mathcal{T},t) + a_{01}(x,\mathcal{T},t) \\ \quad + \alpha_{12} \displaystyle\int_\Omega \int_{SPD(2,\mathbb{R})} w_1(x,\mathcal{T},y,\mathcal{T}')\, \sigma_1\,[s_1(a_1(x,\mathcal{T},t), \\ \qquad\qquad a_1(y,\mathcal{T}',t))]\; dy\, d\mathcal{T}' \\ \dfrac{\partial a_2(x,c,t)}{\partial t} = -\alpha_{21}\, a_2(x,c,t) + a_{02}(x,c,t) \\ \quad + \alpha_{22} \displaystyle\int_\Omega \int_{\mathfrak{C}_{opp}} w_2(x,c,y,c')\, \sigma_2\,[s_2(a_2(x,c,t), \\ \qquad\qquad a_2(y,c',t))]\; dy\, dc', \end{cases}$$

with a_1 denoting the activity of the neurons responding to edges and textures, and a_2 the activity of the neurons responding to colors. The weights w_1 and w_2 are given by (18) and (19).

3.2 On the existence and uniqueness of the solutions of model (17)

In this section, we assume that the vector bundle in the model (17) is of any rank n. In this context, the variational models in the system (20) correspond to particular cases of the variational model (17) for $\gamma_2 = 0$ (first model) or $\gamma_1 = 0$ (second model). Hence, it suffices to study model (17) in what follows. We obtain the following result.

Proposition 2 *The functional in* (17) *is coercive and bounded from below for* $\gamma_2 < 1/2$.

Proof Let us denote by $X = X_1 + X_2$ the functional in (17), where

$$X_1(a) = \frac{1}{2}\|a - a_0\|^2_{L^2(E)} - \frac{\gamma_1}{2}\int_{\Omega^2} w_1(x,y)\|\,\tau_{y,x,\gamma_{y,x}} a(y) - a(x)\|_{h(x)}\, dxdy,$$

$$X_2(a) = -\frac{\gamma_2}{4}\int_{\Omega^2} w_2(x,y)\|\,\tau_{y,x,\gamma_{y,x}} a(y) - a(x)\|^2_{h(x)}\, dxdy.$$

We have

$$\frac{1}{2}\|a - a_0\|^2_{L^2(E)} \geq \frac{1}{2}\|a\|^2_{L^2(E)} - \|a\|_{L^2(E)}\|a_0\|_{L^2(E)} + \frac{1}{2}\|a_0\|^2_{L^2(E)} \tag{21}$$

from Cauchy-Schwarz inequality.
We deduce that

$$X_1(a) \geq \frac{1}{2}\|a\|^2_{L^2(E)} + \frac{1}{2}\|a_0\|^2_{L^2(E)} - \|a\|_{L^2(E)}\|a_0\|_{L^2(E)} \tag{22}$$

if $\gamma_1 \leq 0$ as the second term of X_1 is positive for all $a \in L^2(E)$. It follows from (22) that X_1 is coercive and bounded from below if $\gamma_1 \leq 0$.

Let us now assume that $\gamma_1 > 0$. From the triangle inequality and the fact that the covariant derivative is compatible with h, we have

$$\|\tau_{y,x,\gamma_{y,x}} a(y) - a(x)\|_{h(x)} \leq \|a(x)\|_{h(x)} + \|a(y)\|_{h(y)}.$$

Then, from Fubini's theorem and the fact that w_1 is normalized, we have

$$\int_{\Omega^2} w_1(x,y)\|a(x)\|_{h(x)}dxdy = \int_{\Omega^2} w_1(x,y)\|a(y)\|_{h(y)}dxdy$$

$$= \int_{\Omega} \|a(x)\|_{h(x)}dx.$$

Hence, if $\gamma_1 > 0$, we have

$$-\frac{\gamma_1}{2}\int_{\Omega^2} w_1(x,y)\| \tau_{y,x,\gamma_{y,x}} a(y) - a(x)\|_{h(x)}\, dxdy$$
$$\geq -\gamma_1\|a\|_{L^1(E)},$$

and together with (21) further

$$X_1(a) \geq \frac{1}{2}\|a\|^2_{L^2(E)} + \frac{1}{2}\|a_0\|^2_{L^2(E)} - \|a\|_{L^2(E)}\|a_0\|_{L^2(E)}$$

$$- \gamma_1\|a\|_{L^1(E)}.$$

Then, using Hölder's inequality, we have

$$\|a\|_{L^1(E)} \leq \sqrt{|\Omega|}\,\|a\|_{L^2(E)},$$

which gives

$$X_1(a) \geq \frac{1}{2}\|a\|^2_{L^2(E)} + \frac{1}{2}\|a_0\|^2_{L^2(E)} - \|a\|_{L^2(E)}\|a_0\|_{L^2(E)}$$
$$- \gamma_1\sqrt{|\Omega|}\|a\|_{L^2(E)}. \tag{23}$$

We deduce from (23) that X_1 is coercive and bounded from below for $\gamma_1 > 0$ as well.

If $\gamma_2 \leq 0$, the functional X is coercive and bounded from below as the sum of the coercive and bounded from below functional X_1 and the positive functional X_2.

Finally, let us assume that $\gamma_2 > 0$. Then, inequality (23) together with (15) gives

$$X(a) \geq \frac{1}{2}\|a\|^2_{L^2(E)} + \frac{1}{2}\|a_0\|^2_{L^2(E)} - \|a\|_{L^2(E)}\|a_0\|_{L^2(E)}$$
$$- \gamma_1\sqrt{|\Omega|}\,\|a\|_{L^2(E)} - \gamma_2\|a\|^2_{L^2(E)},$$

which ensures that X is coercive and bounded from below if $\gamma_2 < 1/2$. □

Under the assumption that $\gamma_2 < 1/2$, the functional in (17) is continuous, coercive and bounded from below, which guarantees the existence of a solution of the variational problem (17).
Moreover, assuming in addition that $\gamma_1 \leq 0$, the functional is strictly convex as the sum of a strictly convex functional (see Prop. 1) and a convex functional (the term (4) is convex as the sup of linear functions), which guarantees the uniqueness of the solution of the variational problem (17).

3.3 Numerical solution schemes

In this section, we suggest numerical schemes for finding a minimizer of (17) for $\gamma_2 < \frac{1}{2}$, where we distinguish the cases $\gamma_1 \leq 0$, where the functional is strictly convex and $\gamma_1 > 0$, where the functional is the difference of two convex functions.

3.3.1 The case $\gamma_1 \leq 0$ and $\gamma_2 < 1/2$

Problem (17) is of the form

$$\min_{a\in L^2(E)} \{G(a) + F(Ka)\} \tag{24}$$

where

$$G(a) = -\frac{1}{\gamma_1}\|a - a_0\|^2_{L^2(E)} + \frac{\gamma_2}{2\gamma_1}\|\nabla^{NL}_{w_2} a\|^2_{L^2(\mathrm{pr}_1(E))},$$
$$F(Ka) = \|\nabla^{NL}_{w_1} a\|_{L^1(\mathrm{pr}_1(E))}. \tag{25}$$

Here $K\colon L^2(E) \longrightarrow L^2(\mathrm{pr}_1(E))$ is the nonlocal covariant derivative operator $\nabla^{NL}_{w_1}$. Recall that by Proposition 1 the function G is strictly convex. Let

$$F^*(\eta) = \sup_x\{\langle x, \eta\rangle - F(x)\}$$

denote the (convex) Fenchel conjugate of F. The dual formulation of problem (24) is

$$-\min_{\eta\in L^2(\mathrm{pr}_1(E))}\{G^*(-K^*\eta) + F^*(\eta)\}, \tag{26}$$

i.e., (24) and (26) have the same value. The function F^* is the indicator function of the set

$$\mathcal{H}_1 \colon = \left\{\zeta \in \Gamma(\mathrm{pr}_1(E)), \|\zeta(x,y)\|_{h(x)} \leq 1 \ \ \forall x, y \in \Omega\right\}.$$

Let us now assume that Ω is a discrete grid. In this context, the spaces $L^2(E)$ and $L^2(\mathrm{pr}_1(E))$ are finite-dimensional real vector spaces equipped with an inner product. Moreover, the properties of F, G, K hold in the discrete case well, i.e. the map $K\colon L^2(E) \longrightarrow$

$L^2(\mathrm{pr}_1(E))$ is a continuous linear operator possessing an adjoint K^*, and the maps G, F are proper, convex, lower semi-continuous.

As a consequence, problem (24) can be solved by the primal-dual hybrid algorithm of Chambolle and Pock [14] described in Algorithm 1. In our numerical examples we choose $\theta = 1$. Since F^* is the indicator function of the convex set $\mathcal{H}_1$, the proximal operator of σF^* is given by

$$\mathrm{prox}_{\sigma F^*}(\zeta) = \frac{\zeta}{\max(1, \|\zeta\|_h)}. \tag{27}$$

By definition, the proximal operator of τG is given by

$$\mathrm{prox}_{\tau G}(f) = \underset{v \in L^2(E)}{\arg\min} \frac{1}{2}\|v - f\|^2_{L^2(E)} - \frac{\tau}{\gamma_1}\|v - a_0\|^2_{L^2(E)} + \frac{\tau\gamma_2}{2\gamma_1}\|\nabla^{NL}_{w_2} v\|^2_{L^2(\mathrm{pr}_1(E))}.$$

In particular, assuming that the covariant derivative is flat and denoting by P the moving frame in which the corresponding connection 1-form vanishes, we have

$$\mathrm{prox}_{\tau G}(f) = \mathcal{F}^{-1}\left(\frac{\mathcal{F}\left(P^{-1}a_0 - \frac{\gamma_1}{2\tau}f\right)}{1 - \gamma_2 - \frac{\gamma_1}{2\tau} + \gamma_2\mathcal{F}(w_2)}\right),$$

in the frame P, and the iterative procedure in Algorithm 1 reads

$$a^{n+1} = \mathcal{F}^{-1}\left(\frac{\mathcal{F}\left(P^{-1}a_0 - \frac{\gamma_1}{2\tau}\left(a^n - \tau\nabla^{NL^*}_{w_1}\eta^n\right)\right)}{1 - \gamma_2 - \frac{\gamma_1}{2\tau} + \gamma_2\mathcal{F}(w_2)}\right),$$
$$\overline{a}^n = a^{n+1} + \theta(a^{n+1} - a^n),$$
$$\eta^{n+1} = \frac{\eta^n + \sigma\nabla^{NL}_{w_1}\overline{a}^n}{\max\left(1, \|\eta^n + \sigma\nabla^{NL}_{w_1}\overline{a}^n\|_2\right)}.$$

Finally, denoting by $(\underline{a}, \underline{\eta})$ the limit point of the iteration, the solution of the primal problem (24) in the original frame is $P\underline{a}$.

3.3.2 The case $\gamma_1 > 0$ and $\gamma_2 < 1/2$

Problem (17) is of the form

$$\min_{a \in L^2(E)} \{G(a) - F(K\,a)\}, \tag{28}$$

where

$$G(a) = \frac{1}{\gamma_1}\|a - a_0\|^2_{L^2(E)} - \frac{\gamma_2}{2\gamma_1}\|\nabla^{NL}_{w_2} a\|^2_{L^2(\mathrm{pr}_1(E))}, \tag{29}$$

and $F \circ K$ is defined in (25). Again, by Proposition 1, the function G is strictly convex. Thus, we have to handle a difference of convex function (DC) problem. Due to Toland [38], the dual formulation of problem (28) reads as

$$\min_{\eta \in L^2(\mathrm{pr}_1(E))} \{F^*(\eta) - G^*(K^*\eta)\}. \tag{30}$$

It has the same solution as the primal problem and coincides moreover with the solution of

$$\min_{a \in L^2(E), \eta \in L^2(\mathrm{pr}_1(E))} \Phi(a, \eta) \coloneqq \{G(a) + F^*(\eta) - \langle Ka, \eta\rangle\} \tag{31}$$

see, e.g., [2, Proposition 1]. We call $(\hat{a}, \hat{\eta})$ a critical point of Φ if

$$K^*\hat{\eta} \in \partial G(\hat{a}) \quad \text{and} \quad K\hat{a} \in \partial F^*(\hat{\eta}). \tag{32}$$

Since G and F are proper, convex and lower semi-continuous, this implies by Fenchel duality that

$$\hat{a} \in \partial G^*(K^*\hat{\eta}) \quad \text{and} \quad \hat{\eta} \in \partial F(K\hat{a}),$$

so that actually

$$K\hat{\eta} \in \partial G(\hat{a}) \cap K^*\partial F(K\hat{a})$$

and

$$K\hat{a} \in \partial F^*(\hat{\eta}) \cap K\partial G^*(K^*\hat{\eta}).$$

As in [37], we call a point $\hat{a}$ a critical point of the primal problem (28) if

$$K\hat{\eta} \in \partial G(\hat{a}) \cap K^*\partial F(K\hat{a}) \neq \emptyset,$$

and $\hat{\eta}$ a critical point of the dual problem (30) if

$$K\hat{a} \in \partial F^*(\hat{\eta}) \cap K\partial G^*(K^*\hat{\eta}) \neq \emptyset.$$

Indeed, we see that if $(\hat{a}, \hat{\eta})$ is a critical point of Φ, then $\hat{a}$ is a critical point of the primal problem and $\hat{\eta}$ of the dual problem.

A classical approach to iteratively find local extrema of DC problems was described in [37] and certain accelerations and specifications thereof were suggested later. These algorithms assume that either the convex or the concave part, or both, are evaluated by one of their subgradients. Recently, a primal-dual proximal algorithm for a special DC problem which includes our setting was suggested in [2]. It turns out that this algorithm is a reinvention of the PALM algorithm [12]. For our setting, the algorithm in [2] coincides with Algorithm 2 below without extrapolation step 2, i.e., with $\theta = 0$. In [35] an acceleration of the PALM algorithm, called iPALM was suggested which uses certain extrapolation steps. For another algorithm in this direction see [3]. Our Algorithm 2 also employs an extrapolation step

Algorithm 1 Primal-Dual Algorithm for the Sum of Convex Functions

1: **Initialization:** Choose $\tau, \sigma > 0$ with $\sigma\tau \leq 1/\|K\|_2^2$ and $(a^0, \eta^0) \in L^2(E) \times L^2(\mathrm{pr}_1(E))$, $\theta \in (0, 1]$
2: **Iterations:** For $n = 0, 1, \ldots$ until a stopping criterion is reached

$$\begin{aligned} a^{n+1} &= \mathrm{prox}_{\tau G}(a^n - \tau K^* \eta^n) \\ \overline{a}^n &= a^{n+1} + \theta(a^{n+1} - a^n) \\ \eta^{n+1} &= \mathrm{prox}_{\sigma F^*}(\eta^n + \sigma K\, \overline{a}^n) \end{aligned}$$

Algorithm 2 Primal-Dual Algorithm for the Difference of Convex Functions

1: **Initialization:** Choose $\tau, \sigma > 0$ with $\sigma\tau \leq 4/(\theta\|K\|_2)^2$ and $(a^0, \eta^0) \in L^2(E) \times L^2(\mathrm{pr}_1(E))$, $\theta \geq 0$
2: **Iterations:** For $n = 0, 1, \ldots$ until a stopping criterion is reached

$$\begin{aligned} a^{n+1} &= \mathrm{prox}_{\tau G}(a^n + \tau K^* \eta^n) \\ \overline{a}^n &= a^{n+1} + \theta(a^{n+1} - a^n) \\ \eta^{n+1} &= \mathrm{prox}_{\sigma F^*}(\eta^n + \sigma K\, \overline{a}^n) \end{aligned}$$

similar to the Chambolle-Pock algorithm which cannot be deduced as a special case of iPALM.

In our numerical experiments we choose $\theta = 1$. The proximal operator of σF^* is given by formula (27), and the proximal operator of τG is, by definition,

$$\mathrm{prox}_{\tau G}(f) = \underset{v \in L^2(E)}{\arg\min} \frac{1}{2}\|v - f\|^2_{L^2(E)} + \frac{\tau}{\gamma_1}\|v - a_0\|^2_{L^2(E)} - \frac{\tau\gamma_2}{2\gamma_1}\|\nabla^{NL}_{w_2} v\|^2_{L^2(\mathrm{pr}_1(E))}.$$

Assuming that the covariant derivative is flat and denoting by P the moving frame in which the corresponding connection 1-form vanishes, we have

$$\mathrm{prox}_{\tau G}(f) = \mathcal{F}^{-1}\left(\frac{\mathcal{F}\left(P^{-1}a_0 + \frac{\gamma_1}{2\tau} f\right)}{1 - \gamma_2 + \frac{\gamma_1}{2\tau} + \gamma_2 \mathcal{F}(w_2)}\right).$$

in the frame P, and the iterative procedure in Algorithm 2 reads

$$a^{n+1} = \mathcal{F}^{-1}\left(\frac{\mathcal{F}\left(P^{-1}a_0 + \frac{\gamma_1}{2\tau}\left(a^n + \tau \nabla^{NL^*}_{w_1} \eta^n\right)\right)}{1 - \gamma_2 + \frac{\gamma_1}{2\tau} + \gamma_2 \mathcal{F}(w_2)}\right),$$

$$\begin{aligned} \overline{a}^n &= a^{n+1} + \theta(a^{n+1} - a^n), \\ \eta^{n+1} &= \frac{\eta^n + \sigma \nabla^{NL}_{w_1} \overline{a}^n}{\max\left(1, \|\eta^n + \sigma \nabla^{NL}_{w_1} \overline{a}^n\|_2\right)}. \end{aligned}$$

The following convergence result is based on the fact that the functions G and F^* satisfy the Kurdyka-Łojasiewicz (KŁ) property.

Theorem 1 *Let Φ be defined by* (31) *with G, F and K as in* (29) *and* (25). *Choose $\theta > 0$ and $\tau, \sigma > 0$ with $\sigma\tau < 4/(\theta\|K\|_2)^2$. Then the sequence $\{(a^n, \eta^n)\}_{n\in\mathbb{N}}$ generated by Algorithm 2 converges to a critical point of Φ, and consequently, the sequence $\{a^n\}_{n\in\mathbb{N}}$ converges to a critical point of the primal problem* (28) *and $\{\eta^n\}_{n\in\mathbb{N}}$ to a critical point of the dual problem* (30).

The proof is given in the Appendix B.

4 Application to color image correction

In this section, we test the models (17) and (20) in different contexts.

In sect. 4.1, we show that the model (20) taken with enhancement parameters ($\gamma_1, \gamma_2 > 0$) improves the quality of landscape photographies obtained with standard digital cameras by increasing their colorfulness and the visibility of their details.

In sect. 4.2, we consider a simplified version of the model (17) by assuming that $\gamma_1 = 0$, which makes the model only processes the contrast of the image. We test the model with the covariant derivative described in sect. 4.2.1 and Appendix A.3., which has the property of making the model preserve some feature (including the hue) of the original image. In sect. 4.2.2, we test the model for contrast reduction ($\gamma_2 < 0$) and the results show that it can improve the visual quality of tone-mapped images which present over-enhanced contrast. In sect. 4.2.3, we test the model for contrast enhancement ($\gamma_2 > 0$), and we show that this covariant derivative prevents the generation of color artifacts.

In sect. 4.3, we test the model (17) for different values of its parameters γ_1, γ_2. In particular, for $\gamma_1, \gamma_2 > 0$, we show that the model greatly enhances image details. Hence, besides improving image quality, this approach

can be used as a preprocessing to improve further tasks like detection.

4.1 Processing the local features of the achromatic component and the contrast of the chromatic component

As mentioned in the introduction, a photography taken with a standard digital camera and displayed on a standard monitor appears as a degraded version of the original scene. In particular, a loss of details and contrast can be observed. Hence, enhancing the local features and contrast of the image is desirable in order to make it perceptually closer to the original scene. It is also well-known that humans have the tendency to prefer enhanced image features.

However, the enhancement model must be carefully selected. Indeed, in many cases, the range of values of the achromatic component (lightness, luminance, etc) of the image is widely spread, with the darkest colors of the scene appearing black and the brightest ones appearing white. As a consequence, enhancing image features can increase the areas of black and white regions of the image, which can provide undesirable results. As an example see the top-right image in Fig. 2, where the model (17) has been applied to the top-left image with $h = \mathbb{I}_3$ and $\omega = 0$ in the $L^*a^*b^*$ frame and with the parameters $\gamma_1 = 0, \gamma_2 = 0.45$. Even if the details of the original image are more visible, the increase of black and white regions makes the result undesirable.
On the other hand, the range of values of the chromatic components is usually small. Indeed, one degradation generated during the acquisition process and related to the loss of contrast and details is the loss of colorfulness (saturation, chroma, etc), which makes the image aesthetically unpleasing. In Fig. 2 bottom-left, the model (20) has been applied to the top-left image, with $\gamma_1 = 0, \gamma_2 = 0.55$, and $h_2 = \mathbb{I}_2, \omega_2 = 0$ in the a^*b^* frame, i.e. the model only enhances the contrast of the chrominance components. It gives a much more pleasing image than the original image, with no increase of black and white regions unlike the top-right figure. However, the details of the original image have not been enhanced.
Finally, we observe at the bottom right image that the desirable result can be achieved by enhancing both the local features of the achromatic component and the contrast of the chromatic components. Indeed, we obtain a colorful image with visible details, and there is no increase of black and white areas. To achieve this result, we apply model (20) to the top-left image with $\gamma_1 = 10, \gamma_2 = 0.55$, $h_1 = \mathbb{I}_1, \omega_1 = 0$ in the frame L^*, and $h_2 = \mathbb{I}_2, \omega_2 = 0$ in the a^*b^* frame.

More examples of this approach are shown in Fig. 3, where the original images in the left column are landscape photographies obtained with a standard digital camera. As the original image in Fig. 2, they present a lack of details and colorfulness. Enhancing the contrast of the chrominance components of the original images by applying the model (20) with $\gamma_1 = 0, \gamma_2 = 0.55$, and $h_2 = \mathbb{I}_2, \omega_2 = 0$ in the a^*b^* frame increases the colorfulness of the images but does not increase the visibility of their details (see images in the center column). On the other hand, by applying the model (20) with $\gamma_1 = 16, \gamma_2 = 0.55$, $h_1 = \mathbb{I}_1, \omega_1 = 0$ in the frame L^*, and $h_2 = \mathbb{I}_2, \omega_2 = 0$ in the a^*b^* frame, we obtain colorful images with much more visible details (see images in the right column).

As a conclusion, the model (20) is able to undo some degradations on landscape photographies when obtained with a standard digital camera by making their details more visible and increasing their colorfulness. It gives more pleasing images which are perceptually closer to the real-world scenes.

4.2 A covariant derivative to process images contrast while preserving their aspects

4.2.1 A covariant derivative for image processing inspired by visual perception

Given a color image a_0, let us consider the connection 1-form ω_{a_0} described in Appendix A.3., and given by formula (36) for $\alpha = 0$ and $h = \mathbb{I}_3$ in the $L^*a^*b^*$ frame, which vanishes in moving frames of the form (37).

In this context, and assuming that $\gamma_1 = 0$, the model (17) consists in processing the norm r_0 of a_0

$$r_0 = \|a_0\|_2 = \sqrt{(L^*(a_0))^2 + (a^*(a_0))^2 + (b^*(a_0))^2}, \quad (33)$$

as $P^{-1}a_0 = (r_0, 0, 0)$, and the solution in the frame P is then of the form $(\lambda r_0, 0, 0)$, for some real-valued function λ, which means that the solution in the $L^*a^*b^*$ frame is of the form $\underline{a} = \lambda a_0$. Then, this model has the property of preserving the hue of a_0 as the hue of a color is invariant with respect to homotheties.

Finally, let us point out to that the quantity (33) is perceptually more accurate than the CIE Lightness L^* to describe the brightness of a color in the sense that it encodes one of the properties described by the Helmholtz-Kohlrausch (H-K) effect [18], which states that chromatic colors appear brighter than achromatic colors. Note that it would be perceptually even more accurate to consider also the second property of the H-K

Fig. 2 Landscape photography features enhancement with the models (17) and (20). Top-left: image obtained with a standard camera. Top-right: Enhancement of the contrast of both the chromatic and achromatic components of the top-left image (model (17)). Bottom-left: Enhancement of the contrast of the chromatic components of the top-left image (model (20)). Bottom-right: Enhancement of the contrast of the chromatic components and the local features of the achromatic component of the top-left image (model (20)).

effect, which states that the brightness is hue dependent as well (38), but it would make the contrast processing model lose its hue preserving property. Here we like to mention that there exist other hue preserving color enhancement methods based on variational models as [22,29–31].

4.2.2 Regularizing the contrast of (tone-mapped) high dynamic range images

As mentioned in the introduction, if the tone mapping method is not well adjusted, the features of its output image can look over-enhanced and unrealistic when displayed on a standard monitor. Then, we claim that regularizing the features of the image in an accurate manner can be a good strategy to overcome those problems. In Fig. 4, we test this approach on the two left images, available at `https://unsplash.com/images/stock/hdr`. We observe in these examples that some regions are too dark (see for instance the bottom left corner in the top-left image) and other regions are too bright (see the clouds in the bottom-left image), which suggests that reducing the contrast in these images can improve their visual quality. Then, we apply the model (17) with $\gamma_1 = 0$ and $\gamma_2 = -0.55$ for two different covariant derivatives compatible with the metric $h = \mathbb{I}_3$ in the $L^*a^*b^*$ frame: the one given by $\omega = 0$ (see Fig. 4(center)), and the one described in sect. 4.2.1 (see Fig. 4(right)). We observe in both cases that the defects produced by the tone-mapping are removed or least reduced. However, whereas the first covariant derivative tends to produce results with unsaturated colors, the second one produces results whose colors are more faithful to the colors of the original image (see for instance the sky in the top images).

4.2.3 Enhancing the contrast of low dynamic range images

The camera acquisition process can also affect the hue, i.e. it produces an image whose hue is not necessarily the one of the original scene. Hence, if one wants to reproduce the perception of the real-world scene, it can be desirable that the color correction model modifies the hue of the input image.

However, in some situations, the non preservation of the hue can generate color distortions. For instance,

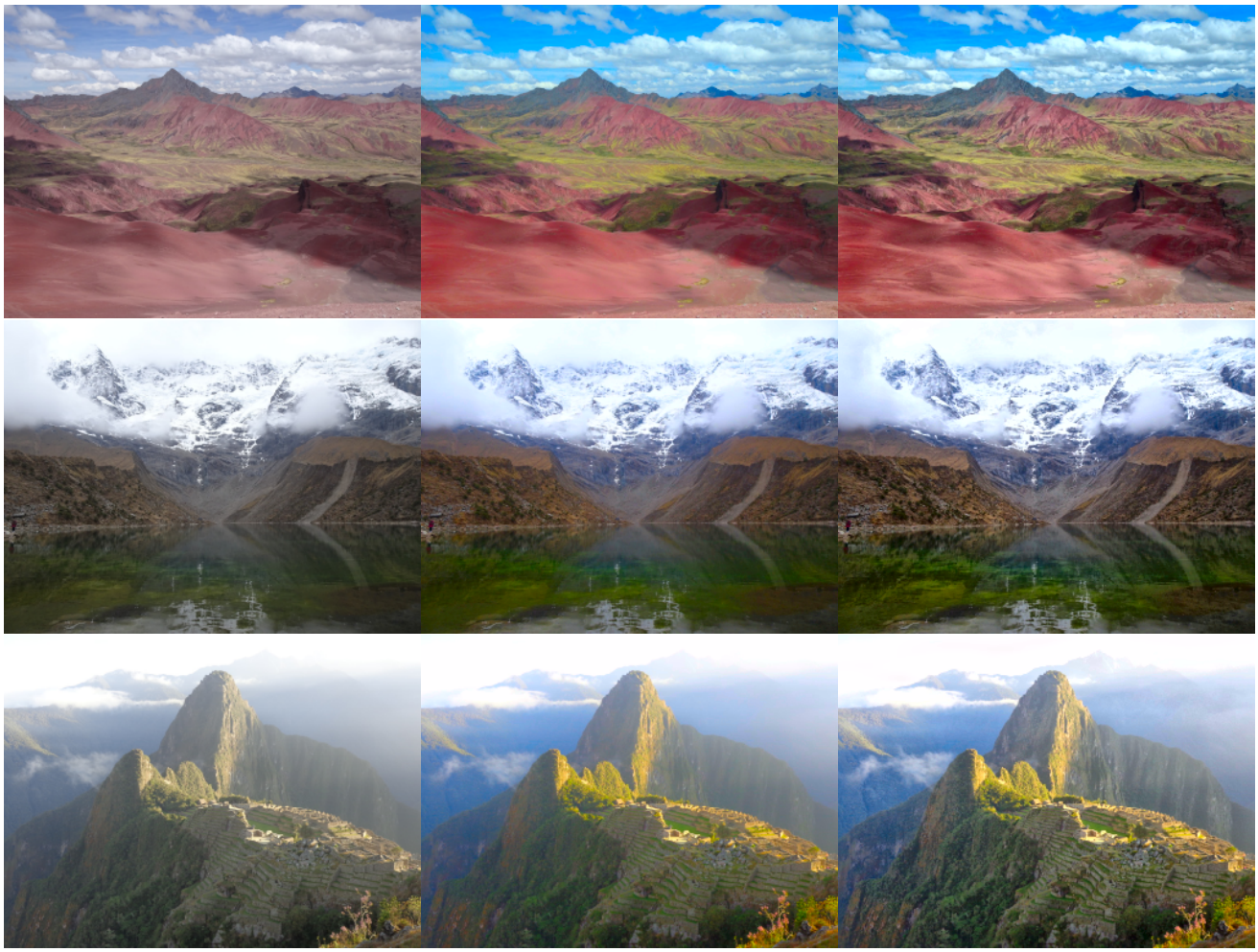

Fig. 3 Landscape photographies features enhancement with the model (20). Left: images obtained with a standard camera. Center: Enhancement of the contrast of the chromatic components of the left images. Right: Enhancement of both the contrast of the chromatic components and the local features of the achromatic component of the left images.

by applying the model (17) to the images in Fig. 5(a-e), with $\gamma_1 = 0, \gamma_2 = 0.45$ and $h = \mathbb{I}_3, \omega = 0$ in the $L^*a^*b^*$ frame, which enhances their contrast, we observe that the white tablecloth in Fig. 5(e) turns blue (see Fig. 5(g)) and the orange geometric figure in the cubist style portrait at the top-right part in Fig. 5(a) turns red (see Fig. 5(c)).

It is also important to avoid over-enhancement when enhancing image features. In this example, the value 0.45 of the enhancement parameter γ_2 provides a good lightness contrast but the chroma contrast is too high. Indeed, we can observe in Fig. 5(c) over-saturated blue at the bottom-right corner and under-saturated colors at the bottom-center dark region where the words are barely readable.

There exist straightforward strategies to address the two issues aforementioned. The first one consists in processing only the lightness of the image and preserve its chrominance, but it can produce results with a lack of colorfulness and consequently unpleasing images (see Fig. 5 (b-f) where the model (17) has been only applied to the L^* component of the images in Fig. 5(a-e) with parameters $\gamma_1 = 0, \gamma_2 = 0.45$ and $\omega \equiv 0$ in the frame L^*). The second one consists in converting the color space $L^*a^*b^*$ into the lightness-chroma-hue color space, and applying this contrast enhancement model to both the lightness and the chroma of the image. However, it would require the tuning of two contrast enhancement parameters γ_2, one per component.

Considering the model (17) with the covariant derivative described in sect. 4.2.1 enables to overcome these issues. Indeed, it has been shown in sect. 4.2.1 that it preserves the hue of the original image and process simultaneously the lightness and the chroma components. We applied this model to the images in Fig. 5(a-e) with parameters $\gamma_1 = 0, \gamma_2 = 0.45$, and the results are shown in Fig. 5 (d-h). As we can see, the contrasts of both the lightness and chroma components are high

Fig. 4 Contrast reduction of tone-mapped high dynamic range images. Left: output of a tone-mapping method. Center: Contrast reduction model (17) induced by the connection 1-form vanishing in the $L^*a^*b^*$ frame. Right: Contrast reduction model (17) induced by the"perceptual" connection 1-form described in sect. 4.2.1.

but not over-enhanced, and the hue modifications observed in Fig. 5(c-g) have disappeared.

4.3 Simultaneous and independent local features and contrast processing

We test the model (17) with the connection 1-form described in sect. 4.2.1 for $\gamma_1 \in \{-8, 0, 8\}$ and $\gamma_2 \in \{-0.25, 0, 0.25\}$. If $\gamma_1 \neq 0$, we solve the model by the iterative procedure described in sect. 3.3. In the experiments performed in this section, the image size is 256×384, the window size of w_1 is 41×41, and the variance of the Gaussian kernel w_2 is 1000. Finally, the parameters τ, σ of the primal dual numerical schemes are set to 0.1 and the stopping criteria is

$$\sqrt{MSE(a^{n+1}, a^n)} < 0.001.$$

Fig. 6 shows the results of applying the model to the images in Fig. 6(e-n). As expected, we observe that the model does enhance resp. regularize the local structures for $\gamma_1 > 0$ resp. $\gamma_1 < 0$ and the contrast for $\gamma_2 > 0$ resp. $\gamma_2 < 0$, providing some flexibility in the processing of image features that the models in [6],[8] do not. In particular, the results show that enhancing simultaneously the contrast and the local features in Fig. 6(i-r) boosts the enhancement of the local structures in Fig. 6(f-o). Indeed, we observe that the details are much more visible in Fig. 6(i) compared to Fig. 6(f) and in Fig. 6(r) compared to Fig. 6(o). Hence, besides improving image quality, this approach can be used as a preprocessing to improve further tasks like detection.

5 Conclusion

We have introduced two new models for color image correction inspired by visual perception through the notion of perceptual gradient and visual neuroscience through the W-C equations describing the temporal evolution of the activity of neurons in the visual cortex. The proposed models are more elaborated than existing variational models. In particular, they rely on the presence of neurons responding to different attributes in the visual cortex: edges/textures and colors, which enabled us to design image processing models correcting different types of degradations on the images involving the enhancement/regularization of details and contrast. Further work will be devoted to generalize our models by including more visual attributes.

Acknowledgements The authors thank the anonymous reviewers for helpful remarks and suggestions.

Appendix A. Covariant derivatives and visual perception

A.1. Interpretation of the Kernel-Based Retinex model

The original Retinex formulation [28] and the Kernel-Based-Retinex [10] can be formulated as follows: The perceived color at a given pixel of the image results from a weighted averaging of the perceptual difference between the given pixel and the other pixels in the image domain. This suggests that the key object in color

perception is the perceived gradient, and that the accuracy of the estimation of the perceived colors depends on the accuracy of the estimation of the perceived gradient.

More precisely, given an RGB color image $a = (a^1, a^2, a^3)\colon \Omega \subset \mathbb{R}^2 \longrightarrow \mathbb{R}^3$, the perceived image $L = (L^1, L^2, L^3)$ is, according to Kernel-Based Retinex [10], given for $k = 1, 2, 3$, by

$$L^k(x) = \int_{y:a^k(y)\geq a^k(x)} w(x,y)\left[A\log\left(\frac{a^k(x)}{a^k(y)}\right)+1\right]dy + \int_{y:a^k(y)<a^k(x)} w(x,y)\,dy,$$

where w is a Gaussian kernel and A is a constant, which can be rewritten as

$$\begin{aligned} L^k(x) &= \int_{y\in\Omega} w(x,y)\,\zeta\left(\log[a^k(x)] - \log[a^k(y)]\right)dy \\ &= \int_{y\in\Omega} w(x,y)\,\zeta\left(\int_{\gamma_{y,x}} d_{\gamma'_{y,x}(t)}\log(a^k)(\gamma_{y,x}(t))dt\right)dy \end{aligned}$$

for some nonlinear function ζ, and for any path $\gamma_{y,x}$ joining y and x.

Then, it turns out that the quantity $\nabla\log(a^k) = da^k/a^k$ can be interpreted as the perceived gradient of the image according to Weber's law in vision and the quantity

$$\int_{\gamma_{y,x}} d_{\gamma'_{y,x}(t)}\log(a^k)(\gamma_{y,x}(t))dt. \tag{34}$$

as the induced perceptual difference between x and y. Indeed, given an uniform background of intensity $\mathcal{I}$, Weber's law states that the following equality holds

$$\frac{\delta\mathcal{I}}{\mathcal{I}} = c,$$

where $\delta\mathcal{I}$ is the minimum intensity increment of $\mathcal{I}$ to which the human sensitivity distinguish $\mathcal{I}$ and $\mathcal{I}+\delta\mathcal{I}$, and c is a constant. Hence, Weber's law shows that the human sensitivity to an intensity increment depends on the intensity of the background. In particular, it shows that human perception is more sensitive to intensity changes in dark backgrounds than in the bright ones.

A.2. Equivariance property of the perceived gradient

Based on the assumption that the color constancy property of the HVS comes from an invariance property of the perceived gradient with respect to lighting changes and under the identification between these latter and moving frame changes on a vector bundle, Georgiev [21] suggested that a covariant derivative is a good candidate to describe the perceived gradient, due to the invariance of this differential operator with respect to moving frame changes. More precisely, given a Lie group G, a G-associated vector bundle E, and a covariant derivative $\nabla\colon = d + \omega$ on E, we have $\nabla(\mathcal{G}a) = \mathcal{G}\nabla(a)$ for any G-valued moving frame $\mathcal{G}$ and section a of E.

The perceptual difference formula (34) can then be modified by means of a covariant derivative, which gives

$$\int_{\gamma_{y,x}} \nabla_{\gamma'_{y,x}(t)} a(\gamma_{y,x}(t))\,dt = \tau_{y,x,\gamma_{y,x}}a(y) - a(x). \tag{35}$$

Note that the quantity (35) is independent of the path $\gamma_{y,x}$ provided that ∇ is flat.

A.3. A perceptual law derived from a covariant derivative compatible with the metric: Helmholtz-Kohlrausch effect

The Helmholtz-Kohlrausch effect is a color appearance phenomena which states that the brightness of a color depends not only on its achromatic component but also on its chromatic component. More precisely, it states that chromatic colors appear brighter than achromatic colors, and some hues appear brighter than others.

Given a color image $a = (a^1, a^2, a^3)$, let us consider the $\mathrm{SO}(h)$-associated vector bundle $\Omega\times\mathbb{R}^3 \longrightarrow \Omega$, for some metric h, and the connection 1-form ω_a given by

$$\begin{pmatrix} 0 & \frac{a^1da^2 - a^2da^1}{\alpha+\|a\|_h^2} & \frac{a^1da^3 - a^3da^1}{\alpha+\|a\|_h^2} \\ -\frac{a^1da^2 - a^2da^1}{\alpha+\|a\|_h^2} & 0 & \frac{a^2da^3 - a^3da^2}{\alpha+\|a\|_h^2} \\ -\frac{a^1da^3 - a^3da^1}{\alpha+\|a\|_h^2} & -\frac{a^2da^3 - a^3da^2}{\alpha+\|a\|_h^2} & 0 \end{pmatrix}, \tag{36}$$

for $\alpha \geq 0$. Having $\omega \in \Gamma(T^*\Omega\otimes\mathfrak{so}(h))$ makes the corresponding covariant derivative ∇ be compatible with h.

At the limit case $\alpha = 0$, it has been shown in [6] that the corresponding covariant derivative ∇ is flat, from which follows the existence of a moving frame P in which ω_a vanishes. More precisely, denoting by $(r\sin\theta\cos\varphi, r\sin\theta\sin\varphi, r\cos\theta)$ the spherical coordinates of a, then P is of the form (37).

$$P = \begin{pmatrix} \cos\varphi\sin\theta & -g_{11}\sin\varphi - g_{21}\cos\theta\cos\varphi & -g_{12}\sin\varphi - g_{22}\cos\theta\cos\varphi \\ \sin\varphi\sin\theta & g_{11}\cos\varphi - g_{21}\sin\varphi\cos\theta & g_{12}\cos\varphi - g_{22}\sin\varphi\cos\theta \\ \cos\theta & g_{21}\sin\theta & g_{22}\sin\theta \end{pmatrix} \tag{37}$$

with $g_{11}, g_{12}, g_{21}, g_{22} \in C^1(\Omega)$ satisfying

$$\begin{cases} g_{11}^2 + g_{21}^2 = 1 \\ g_{12}^2 + g_{22}^2 = 1 \\ g_{11}g_{12} + g_{21}g_{22} = 0 \\ g_{11}\,dg_{12} + g_{21}\,dg_{22} = \cos\theta\,d\varphi. \end{cases}$$

Then, assuming that the coordinates (a^1, a^2, a^3) correspond to the CIE $L^*a^*b^*$ components of a, and the metric h is given by

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \xi^2 & 0 \\ 0 & 0 & \xi^2 \end{pmatrix}$$

in this frame, where

$$\xi = (2.5 - 0.0025L^*)\left(0.116\left|\sin\left(\frac{H^* - 90}{2}\right)\right| + 0.085\right),$$

for H^* denoting the hue component. The quantity

$$r = \|a\|_h = \sqrt{(L^*(a))^2 + \xi^2\left((a^*(a))^2 + (b^*(a))^2\right)} \tag{38}$$

can be identified to the brightness defined by Fairchild and Pirrotta [18], and which takes into account the Helmholtz-Kohlrausch effect.
As a consequence, the perceptual difference (35), which is given here by

$$P^{-1}(y)\,a(y) - P^{-1}(x)\,a(x)$$

in the frame P can be identified to the brightness difference between $a(y)$ and $a(x)$, as $P^{-1}a = (r, 0, 0)$.

B. Proof of Theorem 1

Semi-algebraic functions are functions which graph is a semi-algebraic set, i.e., a finite intersections of polynomial sets $P(x) = 0$ and $Q(x) < 0$ and finite unions thereof. Recall that semi-algebraic functions fulfill the Kurdyka-Łojasiewicz (KŁ) property, see [12]. In particular, the function Φ in (31) satisfies the KŁ property. Let us assume that after discretization $\Phi : \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \to \mathbb{R} \cup \{+\infty\}$.

The following theorem can be found in [1, Theorem 2.9].

Theorem 2 *Let $f\colon \mathbb{R}^d \to \mathbb{R} \cup \{\infty\}$ fulfill the KŁ property. Let $\{x^n\}_{n\in\mathbb{N}}$ be a sequence which fulfills the following conditions:*

i) There exists $K_1 > 0$ such that $f(x^{n+1}) - f(x^n) \leq -K_1\|x^{n+1} - x^n\|^2$ for every $n \in \mathbb{N}$.
ii) There exists $K_2 > 0$ such that for every $n \in \mathbb{N}$ there exists $w_{n+1} \in \partial_L f(x^{n+1})$ with $\|w_{n+1}\| \leq K_2\|x^{n+1} - x^n\|$, where $\partial_L f$ denotes the Fréchet limiting subdifferential of f.
iii) There exists a convergent subsequence $\{x^{n_j}\}_{j\in\mathbb{N}}$ with limit $\hat{x}$ and $f(x^{n_j}) \to f(\hat{x})$.

Then the whole sequence $\{x^n\}_{n\in\mathbb{N}}$ converges to point $\hat{x}$ which fulfills $0 \in \partial_L f(\hat{x})$.

In the following proof we can partially use arguments from [2].

Proof of Theorem 1: We show that the function $\Phi : \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \to \mathbb{R} \cup \{\infty\}$ in (31) and the sequence $\{(a^n, \eta^n)\}_{n\in\mathbb{N}}$ generated by Algorithm 1 fulfills the properties i)-iii) of Theorem 2.
i) By the variational inequality of the proximal operator it holds

$$\begin{aligned} G(a^{n+1}) - G(a) &\leq \frac{1}{\tau}\langle a^n + \tau K^*\eta^n - a^{n+1}, a^{n+1} - a\rangle \\ &= \frac{1}{\tau}\langle a^n - a^{n+1}, a^{n+1} - a\rangle \\ &\quad + \langle K^*\eta^n, a^{n+1} - a\rangle, \\ F^*(\eta^{n+1}) - F^*(\eta) &\leq \frac{1}{\sigma}\langle \eta^n + (1+\theta)\sigma K a^{n+1} - \theta\sigma K a^n \\ &\quad - \eta^{n+1}, \eta^{n+1} - \eta\rangle \\ &= \frac{1}{\sigma}\langle \eta^n - \eta^{n+1}, \eta^{n+1} - \eta\rangle \\ &\quad + \langle (1+\theta)Ka^{n+1} - \theta K a^n, \eta^{n+1} - \eta\rangle. \end{aligned}$$

Choosing $a = a^n$ and $\eta = \eta^n$, this yields

$$\begin{aligned}\Phi(a^{n+1},\eta^n) - \Phi(a^n,\eta^n) &= G(a^{n+1}) - G(a^n)\\ &\quad + \langle K^*\eta^n, a^n - a^{n+1}\rangle\\ &\leq -\frac{1}{\tau}\|a^n - a^{n+1}\|_2^2.\end{aligned}$$

Since $2uv \leq \alpha u^2 + \frac{1}{\alpha}v^2$ for an arbitrary $\alpha > 0$, we get

$$\begin{aligned}&\Phi(a^{n+1},\eta^{n+1}) - \Phi(a^{n+1},\eta^n)\\ &= F^*(\eta^{n+1}) - F^*(\eta^n) + \langle \eta^n - \eta^{n+1}, Ka^{n+1}\rangle\\ &\leq -\frac{1}{\sigma}\|\eta^n - \eta^{n+1}\|_2^2 + \theta\langle Ka^{n+1} - Ka^n, \eta^{n+1} - \eta^n\rangle\\ &\quad - \frac{1}{\tau}\|a^n - a^{n+1}\|_2^2\\ &\leq -\frac{1}{\sigma}\|\eta^n - \eta^{n+1}\|_2^2 + \theta\|K\|_2\|a^n - a^{n+1}\|_2\|\eta^n - \eta^{n+1}\|_2\\ &\leq \Big(-\frac{1}{\sigma} + \frac{\theta\|K\|_2}{2\alpha}\Big)\|\eta^n - \eta^{n+1}\|_2^2 + \frac{\theta\|K\|_2\alpha}{2}\|a^n - a^{n+1}\|_2^2.\end{aligned}$$

Adding the two inequalities gives

$$\begin{aligned}&\Phi(a^{n+1},\eta^{n+1}) - \Phi(a^n,\eta^n) \qquad (39)\\ &\leq \underbrace{\Big(-\frac{1}{\sigma} + \frac{\theta\|K\|_2}{2\alpha}\Big)}_{c_1}\|\eta^n - \eta^{n+1}\|_2^2\\ &\quad + \underbrace{\Big(-\frac{1}{\tau} + \frac{\theta\|K\|_2\alpha}{2}\Big)}_{c_2}\|a^n - a^{n+1}\|_2^2.\end{aligned}$$

Setting $\alpha := \sqrt{\sigma/\tau}$ we see by $\sigma\tau \leq 4/(\theta\|K\|_2)^2$ that both $c_1 < 0$ and $c_2 < 0$. Thus, i) in Theorem 2 is fulfilled with $K_1 := \max\{-c_1, -c_2\}$. In particular, we obtain

$$\Phi(a^{n+1},\eta^{n+1}) \leq \Phi(a^n,\eta^n). \qquad (40)$$

Further, summing up (39) for $n = 0, ..., N-1$ we get

$$\begin{aligned}\Phi(a^N,\eta^N) - \Phi(a^0,\eta^0) \leq& c_1\sum_{n=0}^{N-1}\|\eta^{n+1} - \eta^n\|_2^2\\ &+ c_2\sum_{n=0}^{N-1}\|a^{n+1} - a^n\|_2^2.\end{aligned}$$

Now the facts, that the left hand side is bounded from below by $\inf_{a,\eta}\Phi(a,\eta) - \Phi(a^0,\eta^0)$ and that $c_j < 0$, $j = 1, 2$ yields

$$\sum_{n=0}^{N-1}\|a^{n+1} - a^n\|_2^2 < \infty \quad \text{and} \quad \sum_{n=0}^{N-1}\|\eta^{n+1} - \eta^n\|_2^2 < \infty. \qquad (41)$$

ii) The construction of the iterates in the algorithm fulfill

$$\begin{aligned}\frac{a^{n-1} - a^n}{\tau} + K^*\eta^{n-1} &\in \partial G(a^n),\\ \frac{\eta^{n-1} - \eta^n}{\sigma} + (1+\theta)Ka^n - \theta Ka^{n-1} &\in \partial F^*(\eta^n).\end{aligned}$$

Consider the function $\tilde{\Phi}(a^n,\eta^n)\colon \mathbb{R}^{d_1}\times\mathbb{R}^{d_2} \to \mathbb{R}\cup\{+\infty\}$, $\tilde{\Phi}(a,\eta) = G(a) + F^*(\eta)$. By the calculus of the convex subdifferential and [34, Proposition 8.12] we get

$$\partial_L\tilde{\Phi}(a^n,\eta^n) = \partial G(a^n)\times\partial F^*(\eta^n).$$

By [34, Exercise 8.8] it holds

$$\begin{aligned}\partial_L\Phi(a^n,\eta^n) &= \partial_L\tilde{\Phi}(a^n,\eta^n) - (K^*\eta^n, Ka^n) \qquad (42)\\ &= (\partial G(a^n) - K^*\eta^n)\times(\partial F^*(\eta^n) - Ka^n).\end{aligned}$$

Thus we have

$$\begin{pmatrix} w_1^n\\ w_2^n\end{pmatrix} := \begin{pmatrix}\frac{a^{n-1}-a^n}{\tau} + K^*(\eta^{n-1} - \eta^n)\\ \frac{\eta^{n-1}-\eta^n}{\sigma} + \theta K(a^n - a^{n-1})\end{pmatrix} \in \partial_L\Phi(a^n,\eta^n).$$

With $(u+v)^2 \leq 2(u^2+v^2)$ and $w^n = (w_1^n, w_2^n)^{\mathrm{T}}$ we obtain

$$\begin{aligned}\|w^n\|^2 \leq& \Big(\frac{2}{\tau^2} + \theta^2\|K\|_2^2\Big)\|a^n - a^{n-1}\|_2^2\\ &+ \Big(\frac{2}{\sigma^2} + \|K\|_2^2\Big)\|\eta^n - \eta^{n-1}\|_2^2,\end{aligned}$$

Hence, property ii) in Theorem 2 is fulfilled with $K_2 := \big(\max\{\frac{2}{\tau^2} + \theta^2\|K\|_2^2,\ \frac{2}{\sigma^2} + \|K\|_2^2\}\big)^{\frac{1}{2}}$.

iii) Since F^* is the indicator function of a compact set $\mathcal{H}_1$, G is coercive and strictly convex and K is a linear operator, we have that the level sets of Φ at (a^0,η^0) given by $\{(a,\eta)\in\mathbb{R}^{d_1}\times\mathbb{R}^{d_2} : \Phi(a,\eta) \leq \Phi(a^0,\eta^0)\}$ are bounded from below. By (40), every infinite sequence $\{(a^n,\eta^n)\}_{n\in\mathbb{N}}$ generated by the algorithm is bounded such that there exists a convergent subsequence $\{(a^{n_j},\eta^{n_j})\}_{j\in\mathbb{N}}$ with limit $(\hat{a},\hat{\eta})$. By Lemma 1, we know that $(\hat{a},\hat{\eta})$ is a critical point of Φ. By (42) and (32) any critical point $(\hat{a},\hat{\eta})$ of Φ fulfills $0 \in \partial_L\Phi(\hat{a},\hat{\eta})$ and conversely. This finishes the proof. □

Lemma 1 *Let $\{(a^n,\eta^n)\}_{n\in\mathbb{N}}$ be the sequence generated by Algorithm 1. Then we have the following:*

i) Any cluster point of $\{(a^n,\eta^n)\}_{n\in\mathbb{N}}$ is a critical point of Φ.

ii) (a^n,η^n) critical point of Φ if and only if $(a^n,\eta^n) = (a^{n+1},\eta^{n+1})$ if and only if $\Phi(a^n,\eta^n) = \Phi(a^{n+1},\eta^{n+1})$.

Proof i) Let $(\hat{a},\hat{\eta})$ be a cluster point and $\{(a^{n_j},\eta^{n_j})\}_{j\in\mathbb{N}}$ a subsequence converging to $(\hat{a},\hat{\eta})$. By the iteration scheme we have

$$\frac{a^{n_j}-a^{n_j+1}}{\tau}+K^*\eta^{n_j}\in\partial G(a^{n_j+1}),$$
$$\frac{\eta^{n_j}-\eta^{n_j+1}}{\sigma}+K\left((1+\theta)a^{n_j+1}-\theta a^{n_j}\right)\in\partial F^*(\eta^{n_j+1}).$$

By (41), the first summands in both expressions tend to zero as $j\to\infty$. Using the closeness of the graphs of ∂G and ∂F^* and passing to the limit, we get $K^*\hat{\eta}\in\partial G(\hat{a})$ and $K\hat{a}\in\partial F^*(\hat{\eta})$. Hence $(\hat{a},\hat{\eta})$ is a critical point of Φ.

ii) The first two statements in ii) are equivalent by the following reason: Let $(a^n,\eta^n)=(a^{n+1},\eta^{n+1})$. Then the iteration scheme implies

$$K^*\eta^n\in\partial G(a^n)$$

and

$$K((1+\theta)a^n-\theta a^n)=Ka^n\in\partial F^*(\eta^n),$$

so that (a^n,η^n) is a critical point of Φ. Conversely, if (a^n,η^n) is a critical point of Φ, then if fulfills the above relations and together with the uniqueness of the proximum we conclude from the iteration scheme that $(a^n,\eta^n)=(a^{n+1},\eta^{n+1})$.

The second two statements in ii) are equivalent by (39). This proves the second assertion.

References

1. Attouch, H., Bolte, J., Svaiter, B.F.: Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward-backward splitting, and regularized Gauss-Seidel methods. Math. Program. **137**(1-2, Ser. A), 91-129 (2013).
2. Banert, S., Bot, R.I.: A general double-proximal gradient algorithm for DC programming. Math. Program. **178**(1-2, Ser. A), 301-326 (2019).
3. Bot, R.I., Csetnel, E.R., Nguyen, D.-K.: A proximal minimization algorithm for structured nonconvex and nonsmooth problems. SIAM J. Optim. **29**(2), 1300-1328 (2019).
4. Batard, T., Sochen, N.: A class of generalized Laplacians devoted to multi-channel image processing. J. Math. Imag. Vis. **48**(3), 517-543 (2014)
5. Batard, T., Bertalmío, M.: A class of nonlocal variational problems on a vector bundle for color image local contrast reduction/enhancement. Geometry, Imag. Comput. **2**(3), 187-236 (2015)
6. Batard, T., Bertalmío, M.: A geometric model of brightness perception and its application to color images correction. J. Math. Imaging Vision **60**(6), 849-881 (2018).
7. Batard, T., Ramon, E., Steidl, G., Bertalmío, M.: A connection between image processing and artificial neural networks layers through a geometric model of visual perception. In: Scale Space and Variational Methods in Computer Vision, Lecture Notes in Computer Science, 11603, Springer, 459-471 (2019).
8. Bertalmío, M., Caselles, V., Provenzi, E., Rizzi, A.: Perceptual color correction through variational techniques. IEEE Trans. Im. Processing **16**(4), 1058-1072 (2007).
9. Bertalmío, M., Cowan, J.D.: Implementing the Retinex algorithm with Wilson-Cowan equations. J. Physiol.-Paris, **103**(1-2) , 69-72 (2009).
10. Bertalmío, M., Caselles, V., Provenzi, E.: Issues about Retinex theory and contrast enhancement. Int. J. Comput. Vis. **83**(1), 101-119 (2009).
11. Bertalmío, M.: Image Processing for Cinema. Chapman & Hall/CRC (2014).
12. Bolte, J., Sabach, S., Teboulle, M.: Proximal alternating linearized minimization for nonconvex and nonsmooth problems. Math. Program. **146**(1-2, Ser. A), 459-494 (2014).
13. Bressloff, P.C., Cowan, J.D.: An amplitude equation approach to contextual effects in visual cortex. Neural Comput. **14**(3), 493-525 (2002).
14. Chambolle, A., Pock, T.: An introduction to continuous optimization for imaging. Acta Numer. **25**, 161-319 (2016).
15. Chossat, P., Faugeras, O.: Hyperbolic planforms in relation to visual edges and textures perception. PLOS Comput. Biol. **5**(12), 1-16 (2009).
16. Cowan, J.D., Bressloff, P.C.: Visual cortex and the Retinex algorithm. Proceedings of SPIE, Volume 4662, Human Vision and Electronic Imaging VII (2002).
17. Cyriac, P., Batard, T., Bertalmío, M.: A non local variational formulation for the improvement of tone mapped images. SIAM J. Imaging Sci. **7**(4), 2340-2363 (2014).
18. Fairchild, M.D., Pirrotta, E.: Predicting the lightness of chromatic objects colors using CIELAB. Color Research and Applications **16**(6), 385-393 (1991).
19. Ferradans, S., Bertalmío, M., Provenzi, E., Caselles, V.: An analysis of visual adaptation and contrast perception for tone mapping. IEEE Trans. Pattern Anal. Mach. Intell. **33**(10), 2002-2012 (2011).
20. Förstner, W., Gülch, E.: A fast operator for detection and precise location of distinct points, corners and centres of circular features. In: Proc. ISPRS Intercom-mission Conference on Fast Processing of Photogrammetric Data, 281-305 (1987).
21. Georgiev, T.: Relighting, Retinex theory, and perceived gradients. Proceedings of Mirage (2005).
22. Getreuer, P.: Automatic color enhancement (ACE) and its fast implementation. IPOL J. Image Process. On Line **2** , 266-277 (2012).
23. Hubel, D., Wiesel, T.N.: Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. J. Physiol. **160**(1), 106-154.2 (1962).
24. Hubel D.H.: Eye, Brain and Vision. Scientific American Library, W.H. Freeman & Co (1988).
25. Hurvich, L.M., Jameson, D.: An opponent-process theory of color vision. Psychol. Rev. **64**(6), 384-404 (1957).
26. Johnson, E.N., Hawken, M.J., Sharpley, R.: The orientation selectivity of color-responsive neurons in macaque V1. J. Neurosci. **28**(32), 8096-8106 (2008).
27. Land, E., McCann, J.J: Lightness and Retinex theory. J. Optical Soc. of Am. **61**(1), 1-11 (1971).
28. Land, E.: The Retinex theory of color vision. Scientific American, 237, 108-128 (1977).
29. Nikolova, M., Steidl, G.: Fast hue and range preserving histogram specification: theory and new algorithms for color image enhancement. IEEE Trans. Image Process. **23**(9), 4087-4100 (2014).

30. Pierre, F., Aujol, J.-F., Bugeau, A., Ta, V.-R.: Luminance-hue specification in the RGB space. In: Scale Space and Variational Methods in Computer Vision, Lecture Notes in Computer Science, 9087 , Springer, 413-424 (2015).
31. Pierre, F., Aujol, J.-F., Bugeau, A., Steidl, G., Ta, V.-T.: Variational contrast enhancement of gray-Scale and RGB images. J. Math. Imag. Vis. **57** (1), 99-116 (2017).
32. Provenzi, E., De Carli, L., Rizzi, A., Marini, D.: Mathematical definition and analysis of the Retinex algorithm. J. Opt. Soc. Am. A, **22**(12), 2613-2621 (2005).
33. Reinhard, E., Ward, G., Pattanaik, S.N., Debevec, P.E., Heidrich, W., Myszkowski, K.: High Dynamic Range Imaging: Acquisition, Display, and Image-Based Lighting. Morgan Kaufmann, San Francisco (2010).
34. Rockafellar, R.T., Wets, R.J.-B.: Variational Analysis, volume 317 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin (1998).
35. Sabach, S., Pock, T.: Inertial proximal alternating linearized minimization (iPALM) for nonconvex and nonsmooth problems. SIAM J. Imag. Sci. **9**(4), 1756-1787 (2016).
36. Song, A., Faugeras, O., Veltz, R.: A neural field model for color perception unifying assimilation and contrast. PLoS Comput. Biol. **15**(6), e1007050 (2019).
37. Tao, P.D., An, L.T.H.: Convex analysis approach to d.c. programming: theory, algorithms, and applications. Acta Math. Vietnam. **22**(1), 289-355 (1997).
38. Toland, J.F.: A duality principle for non-convex optimisation and the calculus of variations. Archiv. Rational Mech. Analysis **71**(1), 41-61 (1979).
39. Wilson H.R., Cowan, J.D.: Excitatory and inhibitory interactions in localized populations of model neurons. Biophys. J., **12**(1), 1-24 (1972).
40. Wilson H.R., Cowan, J.D.: A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. Biol. Cybernet. **13**(2), 55-80 (1973).
41. Yeonan-Kim, J., Bertalmío, M.: Analysis of retinal and cortical components of Retinex algorithms. J. Electron. Imag. **26**(3), 031208 (2017).

(a) Input image (b) Chrominance preserving method

(c) Channelwise method (d) Moving frame method

(e) Input image (f) Chrominance preserving method

(g) Channelwise method (h) Moving frame method

Fig. 5 Comparison of three contrast enhancement methods: channelwise method (model (17) with $\gamma_1 = 0$ and $\omega = 0$ in the $L^*a^*b^*$ frame)), chrominance preserving method (model (17) applied to the L^* component with $\gamma_1 = 0$, $h = \mathbb{I}_1$ and $\omega = 0$ in the L^* frame), and moving frame method (model (17) with $\gamma_1 = 0$ and with the connection 1-form described in sect. 4.2.1).

(a) $\gamma_1 = -8$, $\gamma_2 = -0.25$ (b) $\gamma_1 = 0$, $\gamma_2 = -0.25$ (c) $\gamma_1 = 8$, $\gamma_2 = -0.25$

(d) $\gamma_1 = -8$, $\gamma_2 = 0$ (e) $\gamma_2 = 0$, $\gamma_1 = 0$ (f) $\gamma_1 = 8$, $\gamma_2 = 0$

(g) $\gamma_1 = -8$, $\gamma_2 = 0.25$ (h) $\gamma_1 = 0$, $\gamma_2 = 0.25$ (i) $\gamma_1 = 8$, $\gamma_2 = 0.25$

(j) $\gamma_1 = -8$, $\gamma_2 = -0.35$ (k) $\gamma_1 = 0$, $\gamma_2 = -0.35$ (l) $\gamma_1 = 8$, $\gamma_2 = -0.35$

(m) $\gamma_1 = -8$, $\gamma_2 = 0$ (n) $\gamma_1 = 0$, $\gamma_2 = 0$ (o) $\gamma_1 = 8$, $\gamma_2 = 0$

(p) $\gamma_1 = -8$, $\gamma_2 = 0.35$ (q) $\gamma_1 = 0$, $\gamma_2 = 0.35$ (r) $\gamma_1 = 8$, $\gamma_2 = 0.35$

Fig. 6 Local structures enhancement ($\gamma_1 > 0$) or regularization ($\gamma_1 < 0$) coupled with contrast enhancement ($\gamma_2 > 0$) or regularization ($\gamma_2 < 0$) given by the model (17).