



Total generalized variation-based Retinex image decomposition

Chunxue Wang¹ · Huayan Zhang² · Ligang Liu³

© Springer-Verlag GmbH Germany, part of Springer Nature 2020

Abstract

Human visual system (HVS) can perceive color under varying illumination conditions, and Retinex theory is precisely aimed to simulate and explain how the HVS perceives reflectance regardless of different illumination conditions. In this paper, we introduce a reflectance and illumination decomposition model for the Retinex problem via total generalized variation regularization and H^1 decomposition. The total generalized variation regularization ameliorates the staircasing artifacts that appear in the reflectance component of existing total variation-based models and H^1 norm guarantees smoother illumination. We analyze the existence and uniqueness of the proposed model and employ an alternating minimization scheme based on split Bregman iteration. We present numerous numerical experiments on both grayscale and color images to make comparisons with several state-of-the-art methods and demonstrate that our method is comparable both quantitatively and qualitatively.

Keywords Retinex theory · Image decomposition · Total generalized variation regularization · Alternating minimization scheme

1 Introduction

Retinex theory is originally proposed by Land and McCann [34] to deal with compensation for illumination effects in images, whose idea is to simulate how human visual system (HVS) automatically perceives color under varying illumination conditions. An important feature of HSV is our visual system tends to see the same color in a given image regardless of the light. The primary goal of Retinex theory is to decompose a given image I into two components, reflectance R and illumination L , such that

$$I(x) = R(x) \times L(x), \quad (1)$$

at each pixel $x \in \Omega$ where $\Omega \subset \mathbb{R}^2$ is the image domain. Equation (1) exactly explains why region A of Fig. 1a is visually darker than region B, while they have the same intensity value just shown in Fig. 1b. This is because, by Retinex theory, they are in different illumination conditions. Actually,

the region B is in the shadow of the green cylinder so that the illumination of the region A is stronger than that of the region B, i.e., $L(A) > L(B)$; thus, the reflectance of the region A is smaller than that of the region B, i.e., $R(A) < R(B)$.

To simulate the HVS, we need to recover the reflectance R and illumination L from a given image I accurately, and it is a mathematically ill-posed problem since both parts are unknown. Numerous techniques have been developed to implement the Retinex theory in the past 40 years and a first step taken by most algorithms is the conversion to the logarithmic domain by $i = \log(I)$, $r = \log(R)$, $l = \log(L)$, and thereby, $i = r + l$. For example, path-based algorithm [10,30,31,34,45,55] was proposed with the path geometry of piecewise linear, double spirals and Brownian paths. However, these path-based methods contain a lot of parameters, and they have high computational complexity. The recursive algorithms [13,47] extend the path-based algorithms and replace the path computation by a recursive matrix calculation which highly improves the computational efficiency, but the result strongly depends on the number of iterations. The center/surround algorithms [24,25] assume the illumination component changes smoothly and the reflectance component is computed by subtracting a blurred version of the input image, but these methods need too many parameters. Later on, the idea of path following is formulated into PDE-based models [1,21,49,50] and variational models [28,38,42,43,51,72]. The former recovers the reflectance compo-

✉ Chunxue Wang
chunxuewang2019@163.com

¹ Dunhuang Academy, Gansu, China

² The School of Computer Science and Software Engineering, Tiangong University, Tianjin, China

³ The School of Mathematical Sciences, University of Science and Technology of China, Anhui, China

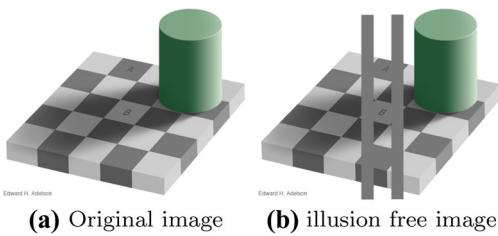


Fig. 1 Adelson's checker shadow illusion. **a** Original image. **b** Illusion free image

ment by solving a Poisson equation with the assumption that the illumination is relative small, and the latter utilizes different regularizations on reflectance and illumination due to their different properties, such as total variation (TV) is introduced to ensure getting piecewise constant reflectance. But it loses too much details of the reflectance component, especially for the high-contrast regions. Recently, deep learning-based image processing methods have been proposed in the Retinex-based image decomposition field [39,53,60,61,70]. As there is no ground truth of the reflectance and illumination, making difficult to generating and annotating dataset of natural images, these methods train their networks either on synthetic dataset, real-world dataset with sparse annotations or image pairs/sequences, and perform not well on real-world data.

In this paper, we consider a decomposition model in the variational framework to recover R and L simultaneously. In particular, we propose to use total generalized variation (TGV) as a regularization term of the reflectance component to capture more details, which has been successfully applied in image restoration [27,29,63], and extract smoother illumination component with H^1 norm. We develop a rather simple, efficient optimization algorithm based on split Bregman iteration, and the keys to our method are recovering richly detailed reflectance and small smooth illumination alternately. We denote our method as TGVH1 reflectance and illumination decomposition model. Compared to the state-of-the-art methods [14,28,38,42,51], we can achieve a more detail-preserved reflectance and smoother illumination and increase the contrast precisely.

The rest of the paper is organized as follows. In Sect. 2, we review several different Retinex algorithms. In Sect. 3, we present the proposed model including theoretical analysis about the existence and uniqueness of the solution. In Sect. 4, we introduce an alternating optimization based on split Bregman iteration to solve the proposed model. In Sect. 5, we present the numerical experiments on visual illusion images, synthetic images, natural color images, medical images and shadowed color images and make comparison with several existing methods. Finally, the conclusion and future work are discussed in Sect. 6.

2 Related work

Generally, image decomposition algorithms can be classified into two categories: intrinsic image decomposition [44] and Retinex-based image decomposition [22]. The former focuses on recovering reflectance close to ground truth as far as possible but losing objects's shape, while the latter aims to restore the reflectance with good shape and visual realism. Here, we mainly work on and review Retinex-based methods. For the past few decades, many interpretations, implementations and improvements of the Retinex algorithm have been studied and basically classified into four categories (see [22,28,36,50] for details): path-based algorithms, recursive algorithms, center/surround algorithms, formulas-based algorithms and supervised deep learning-based algorithms. These seemingly different algorithms are actually very similar since they are all based on the assumption that the illumination component is spatially smooth.

Path-based algorithms The key idea of path-based algorithms was the lightness of each pixel depending on the multiplication of ratios along random walks. Land's original works [30–32,34] used piecewise linear paths and laid out a time-consuming algorithm. Later, Marini and Rizzi [45] implemented the original Retinex algorithm using Brownian motion to approximate each path and greatly improved the effectiveness of the Retinex algorithm. Provenzi et al. [56] presented a detailed mathematical analysis of the original Retinex algorithm to describe the behavior of the algorithm. Base on this work, Provenzi et al. [57] proposed a “random spray Retinex” in which paths were replaced by 2-D pixel sprays, and achieved faster performances than the path-wise one.

Recursive algorithms As a variant of path-based algorithms, recursive algorithms replaced the path computation by a recursive matrix comparison where the comparison between pixels could be calculated simultaneously. Frankle–McCann Retinex [13] used single pixel comparisons with variable separations where a single pixel eventually averages different products from all other pixels data. McCann99 Retinex [47] created a multiresolution pyramid from the input by averaging image, and pixel comparisons and lightness estimates are in an alternate progress from high level to low level until the pyramids bottom level. Both implementations were presented in [15], and the parameter determinations were discussed in [9].

Center/surround algorithms The center/surround approach was first proposed by [33] and later improved by Jobson et al. [24] with an easy implementation and better result. Single-scale Retinex (SSR) [25] and multiscale Retinex (MSR) [58] were the representative algorithms of the cen-

ter/surround approach, and both were based on the idea that lightness values could be computed by subtracting a blurred version of the input image from the original image. Because of their simplicity and effectiveness, many improved algorithms [7,8,23,54,68] were proposed. For example, the multiscale Retinex with color restoration (MSRCR) [59] was an improvement in the MSR while there were halo artifacts near strong edges.

Formulas-based algorithms Formulas-based algorithms were proposed to translate the Retinex principles into a physical model solving by a set of equations or an optimization problem. Poisson equation-type Retinex algorithms were developed by solving a Poisson equation under the assumptions that reflectance was a piecewise constant and illumination was spatially smooth [34]. Horn [21] applied the Laplacian to get $\Delta i = \Delta r + \Delta l$ and then used a threshold function clipping the finite parts to get a Poisson equation, namely

$$\Delta r(x) = \tau(\Delta i), \quad (2)$$

where the threshold function is defined by

$$\tau(x) = \begin{cases} x & \text{if } |x| > t \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

The Poisson equation (2) was solved by an iterative scheme of a low-pass filter. Blake [1] introduced an improvement in Horn's method, extracting the discontinuities from the image gradient magnitude instead of the Laplacian with better boundary conditions. Funt et al. [16] improved Horn's method by applying a new curl-correction technique at the thresholded locations. Additional algorithmic improvements [49,50] by Morel constructed a very tight connection with Horn's and Blake's method and yielded an exact and fast implementation using only two fast Fourier transforms.

By the same assumption with Poisson equation-type Retinex algorithms, variational Retinex algorithms have been applied to image decomposition in the past 2 decades. Let $\Omega \subset \mathbb{R}^d$ be a bounded open domain with a compact Lipschitz boundary, the primary variational model was proposed by Kimmel [28], only focusing on the penalty functional of the illumination component l , and the followed $H^1 + L^2$ model is

$$\begin{aligned} \hat{l} = \arg \min_l & \int_{\Omega} \left\{ |\nabla l|^2 + \alpha(l - i)^2 + \beta |\nabla(l - i)|^2 \right\} dx, \\ \text{s.t. } & l \geq i, \langle \nabla l, \mathbf{n} \rangle = 0 \text{ on } \partial\Omega \end{aligned} \quad (4)$$

which was a quadratic programming problem solved by the project normalized steepest descent method. Ma and Osher

[43] proposed total variation (TV) model to recover piecewise constant reflectance component r , namely

$$\hat{r} = \arg \min_r \int_{\Omega} \left(t |\nabla r| + \frac{1}{2} |\nabla r - \nabla i|^2 \right) dx, \quad (5)$$

where parameter $t > 0$. Afterward, Ma [42] improved a L_1 -based model on reflectance r , which reads

$$\hat{r} = \arg \min_r \int_{\Omega} |\nabla r - \delta_t(\nabla i)| dx, \quad (6)$$

where $\delta_t(\nabla i)$ was gradient threshold function with respect to t . Both models were solved effectively by Bregman iteration, but the penalization on the magnitude of image gradient caused the loss of reflection's details and contrast. The same was to optimize reflectance r , Zosso et al. [72] extended the TV-based models to a unified nonlocal formulation of the reflectance, and Li et al. [37] applied an adaptive perceptually inspired variational model to adjust the uneven intensity distribution.

Several decomposition models have been developed by considering both reflectance r and illumination l simultaneously. For instance, Ng and Wang [51] added an L^2 fidelity term and some constraints in the proposed energy functional, namely

$$\min_{\substack{r \leq 0, \\ l \geq i}} \int_{\Omega} \left\{ |\nabla r| + \frac{\alpha}{2} |\nabla l|^2 + \frac{\beta}{2} (i - r - l)^2 + \frac{\mu}{2} l^2 \right\} dx, \quad (7)$$

where the constraint $r \leq 0$ corresponded to the assumption $0 < R < 1$. Recently, decomposition models combined various kinds of techniques upon r or l with formulation (7) aiming to estimate the reflectance with more details. For example, Chang et al. [5] used sparse and redundant representations of the reflectance component replacing of the TV term. Similarly, Wang et al. [67,69] used nonlocal bounded variation (NLBV) and variational model with barrier functionals upon the reflectance, respectively. Liang and Zhang [38] adopted higher-order total variation L^1 (HoTVL1) on the illumination component. Wang et al. [64] constructed a variational Bayesian model with Gibbs distributions for both reflectance and illumination and the gamma distributions for the model's parameters. In addition, a weighted variational model was proposed in [14] to eliminate impact of log-transform and a variational model with hybrid hyper-Laplacian priors was presented in [6]. Guo et al. [20] and Gao et al. [17] only estimate illumination with structure priors. Gu et al. [19] integrate ASR to represent image large-scale structures and SSR to represent image fine-scale textures.

Supervised deep learning-based algorithms With the rapid development of deep neural network, supervised deep

learning-based algorithms have been proposed in the low-light image enhancement. Lore et al. [39] use stacked-sparse denoising autoencoder to learn adaptively enhance and denoise from synthetically darkened and noise-added training data. Shen et al. [60] analyze the property of MSR algorithm in the sense of convolutional neural network (CNN) and proposed MSR-net to enhance the low-light image trained on synthetically high quality and low-light images pairs. Park [53] combine the stacked and convolutional autoencoders on illumination and reflectance, respectively, also trained on synthetic high-contrast and low-light images. However, those synthetic data are quite different from the real world data, so those methods cannot generalize well to real-world images. To overcome above difficulties, Wei et al. [70] adopt a deep end-to-end Retinex-Net including a Decom-Net and an Enhance-Net, learning without ground truth decompositions under the key constraints that the consistent reflectance is shared by paired low/normal-light images and Shi [62] proposed a generative adversarial network (GAN).

3 TGV + H^1 decomposition model

In this section, we first review the TGV regularization in Sect. 3.1 and Hilbert space H^1 , then formulate the optimization for reflectance and illumination decomposition in Sect. 3.3 and show existence and uniqueness of solution for the above problem in Sect. 3.4.

3.1 Background of total generalized variation

The total generalized variation (TGV) was proposed by Bredies et al. [2], as a generalization of the infimal convolution of TV and second-order TV regularizers [4] is written as follows:

$$\begin{aligned} \text{TGV}_{\alpha}^k(u) = & \sup_{v} \left\{ \int_{\Omega} u \text{div}^k(v) dx \mid v \in \mathcal{C}_c^k(\Omega, \text{Sym}^k(\mathbb{R}^d)), \right. \\ & \left. \| \text{div}^j(v) \|_{\infty} \leq \alpha_j, \quad j = 0, \dots, k-1 \right\}, \end{aligned} \quad (8)$$

where $\alpha = (\alpha_0, \alpha_1, \dots, \alpha_{k-1})$ is fixed positive vector, $\text{Sym}^k(\mathbb{R}^d)$ denotes the space of symmetric k -tensors on \mathbb{R}^d as $\text{Sym}^k(\mathbb{R}^d) = \{\xi : \mathbb{R}^d \times \dots \times \mathbb{R}^d \mid \xi \text{ is } k\text{-linear and symmetric}\}$, and $\mathcal{C}_c^k(\Omega, \text{Sym}^k(\mathbb{R}^d))$ is the space of compactly supported symmetric tensor fields. Especially, it is equivalent to the TV regularizer when $k=1$, $\alpha_0=1$.

It has been shown in [2] that TGV_{α}^k can be interpreted as a k -fold infimal convolution by employing the Fenchel–Rockafellar duality formula, as follows:

$$\text{TGV}_{\alpha}^k(u) = \inf_{u_j} \sum_{j=1}^k \alpha_{k-j} \int_{\Omega} |\mathcal{E}(u_{j-1}) - u_j| dx, \quad (9)$$

where $u_0 = u$, $u_k = 0$, $u_j \in \mathcal{C}_c^{k-j}(\Omega, \text{Sym}^j(\mathbb{R}^d))$ for $j = 1, \dots, k$, and \mathcal{E} represents the distributional symmetrized derivative, i.e., $\mathcal{E}(u_{j-1}) = \frac{\nabla u_{j-1} + (\nabla u_{j-1})^T}{2}$. Thus, TGV_{α}^k automatically balances the first to the k th derivatives of u among themselves and reduces the staircasing effects of the TV. Similar to the definition of the bounded variation space (BV) on a non-empty open set $\Omega \subset \mathbb{R}^d$ in [52], the space of functions of bounded generalized variation of order k with α is defined on Ω in [2] is a Banach space, as follows:

$$\text{BVG}_{\alpha}^k(\Omega) = \left\{ u \in L^1(\Omega) \mid \text{TGV}_{\alpha}^k(u) < \infty \right\}, \quad (10)$$

equipped with the norm $\|u\|_{\text{BVG}_{\alpha}^k} = \|u\|_1 + \text{TGV}_{\alpha}^k(u)$. For instance, when $k=2$, TGV_{α}^2 involves the l_1 -norm of $\mathcal{E}(u_0) - u_1 = \nabla u - u_1$ and $\mathcal{E}(u_1)$ via the Radon norm.

In this work, we particularly utilize the second-order TGV denoted by TGV_{α}^2 and corresponding Banach space $\text{BVG}_{\alpha}^2(\Omega)$ and their mainly analytical properties [2,3] to obtain the existence and uniqueness of the solutions of our proposed model in the last subsection of this part.

3.2 Background of Hilbert Space H^1

We first present some preliminaries for Sobolev spaces proposed in [26,46]. Let Ω be an open subset of \mathbb{R}^n with Lipschitz boundary, the Sobolev space $W^{k,p}(\Omega)$ consists of functions $u \in L^p(\Omega)$ such that for every multi-index with $|\alpha| \leq k$, the weak derivative $D^{\alpha}u$ (same with the classical derivative ∇u for smooth function) exists and $D^{\alpha}u \in L^p(\Omega)$. Thus,

$$W^{k,p}(\Omega) = \{u \in L^p(\Omega) : D^{\alpha}u \in L^p(\Omega), |\alpha| \leq k\},$$

equipped with the norm

$$\|u\|_{W^{k,p}(\Omega)} = \left(\sum_{|\alpha| \leq k} \int_{\Omega} |D^{\alpha}u|^p dx \right)^{\frac{1}{p}}, \quad 1 \leq p < \infty.$$

It is important to emphasize that the Sobolev space $W^{k,p}(\Omega)$ is a Banach space. In particular, $H^k(\Omega) = W^{k,2}(\Omega)$ is a Hilbert space with the inner product,

$$\langle u, v \rangle_{H^k(\Omega)} = \sum_{|\alpha| \leq k} \langle D^{\alpha}u, D^{\alpha}v \rangle_{L^2(\Omega)},$$

and the norm $\|u\|_{H^k(\Omega)} = \langle u, u \rangle^{\frac{1}{2}}$. Here, we mainly focus on the $H^1(\Omega)$ and summarize some definitions and main

properties in $H^1(\Omega)$ and write the $\|u\|_2 = \|u\|_{L^2(\Omega)}$ and $\|u\|_1 = \|u\|_{H^1(\Omega)} = (\|u\|_2^2 + \|\nabla u\|_2^2)^{\frac{1}{2}}$ as a convenience.

(Lower semicontinuity) If the sequence $\{u_k\}_{k \in \mathbb{N}}$ satisfies $u_k \rightarrow u$ and $\nabla u_k \rightarrow \nabla u$, then $\|u\|_1 \leq \liminf_{k \rightarrow \infty} \|u_k\|_1$. In particular, for $\liminf_{k \rightarrow \infty} \|u_k\|_1 < \infty$, then $u \in H^1(\Omega)$.

(Compactness) Suppose $\{u_k\}_k$ is bounded in $H^1(\Omega)$, then there exists a subsequence $\{u_{kj}\}_{j \in \mathbb{N}}$ and $u \in H^1(\Omega)$ such that $u_{kj} \rightarrow u$ and $\nabla u_{kj} \rightarrow \nabla u$ as $k \rightarrow \infty$.

(Embedding) If Ω has a Lipschitz boundary and it is connected and $u \in H^1(\Omega)$, then it can be shown that there exists positive constant C such that $\|u\|_1 \leq C\|\nabla u\|_2$.

3.3 Proposed model

Here, we present $\text{TGV}_{\alpha}^2 + H^1$ variational model for reflectance and illumination decomposition, recovering the reflectance component with more details by TGV_{α}^2 and the spatially smooth illumination component penalized by H^1 norm. We also consider this problem on the logarithmic domain like most variational methods, and the energy function $E(r, l)$ is defined as follows

$$\text{TGV}_{\alpha}^2(r) + \int_{\Omega} \left\{ \frac{c_1}{2} |\nabla l|^2 + \frac{c_2}{2} (i - l - r)^2 + \frac{c_3}{2} l^2 \right\} dx, \quad (11)$$

Since the number of unknowns is twice the number of equations, we further consider constraints \mathcal{B}_r and \mathcal{B}_l on both r and l , aiming to recover the best contrast of the inputs. So the associated minimization problem is given as follows

$$(r^*, l^*) = \underset{(r, l) \in \Lambda \cap \mathcal{B}_r \times \mathcal{B}_l}{\operatorname{argmin}} E(r, l) \quad (12)$$

where $\Lambda = \{(r, l) | (r, l) \in \text{BGV}_{\alpha}^2(\Omega) \times H^1(\Omega)\}$. c_1, c_2 and c_3 are positive parameters, the term $\int_{\Omega} (i - l - r)^2 dx$ is used for the fidelity and $\int_{\Omega} l^2 dx$ is the guarantee of the theory proof. Numerically, c_3 is set very small. $\text{TGV}_{\alpha}^2(r)$ is minimized over all gradients of the deformation field $\mathbf{p} = (p_1, p_2)$ on image space Ω , which reads

$$\inf_{\mathbf{p} \in \mathcal{C}_c^2(\Omega, \mathbb{R}^2)} \alpha_1 \int_{\Omega} |\nabla r - \mathbf{p}| dx + \alpha_0 \int_{\Omega} |\mathcal{E}(\mathbf{p})| dx, \quad (13)$$

where α_0 and α_1 are positive parameters balancing between the first- and the second-order derivative of r .

Remark 1 Here, we show the effect of constraints \mathcal{B}_r and \mathcal{B}_l we used on improving the contrast of estimated reflectance. Figure 2 shows a contrast comparison on different constraints by computing the intensity difference between A and B denoted as $D(A, B)$. (a) shows the input gray image with the same intensity of A and B, (b) takes $\mathcal{B}_r = (-\infty, 0]$ and $\mathcal{B}_l = [i, \infty)$ similar with Ng and Wang [51] and (c) takes $\mathcal{B}_r = (0, 255]$ and $\mathcal{B}_l = [-255, 0)$. It can be noticed that

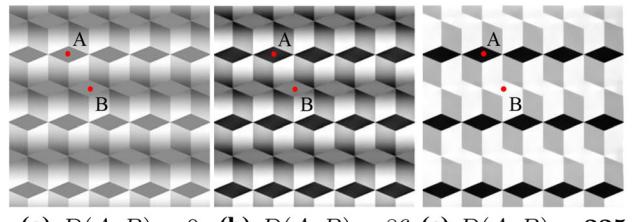


Fig. 2 Contrast comparison with different constraints. **a** shows the input gray image with the intensity of masked A and B 140, $D(A, B) = 0$; **b** shows the result of $\mathcal{B}_r = (-\infty, 0]$ and $\mathcal{B}_l = [i, \infty)$ with the intensity of masked A 73 and B 159, $D(A, B) = 86$; and **c** shows the result of $\mathcal{B}_r = (0, 255]$ and $\mathcal{B}_l = [-255, 0)$ with the intensity of masked A 25 and B 250, $D(A, B) = 225$

constraints in (c) can help to improve the contrast significantly.

3.4 Existence and uniqueness of solution

We give a general existence and uniqueness result for solution of the above minimization problem.

Theorem 1 Suppose $i \in L^{\infty}(\Omega)$ and $c_1, c_2, c_3 > 0$, then problem (12) has a unique solution pair $(r^*, l^*) \in \text{BGV}_{\alpha}^2(\Omega) \times H^1(\Omega)$.

Proof For any $(r, l) \in \Lambda$, the objective functional of (12) is proper and nonnegative (by the property of $\text{BGV}_{\alpha}^2(\Omega)$), i.e., $E(r, l)$ is bounded below. Specifically, it is easy to see that the functional is the correct setting since any constant r and l will yield a finite energy. Hence, we can choose a minimizing sequence $\{r^k, l^k\}_{k \in \mathbb{N}} \subset \Lambda$ for the problem (12), and there exists a constant $M > 0$ such that

$$E(r^k, l^k) \leq M.$$

This yields that

$$\|\nabla l^k\|_2^2 < M, \quad \|l^k\|_2^2 < M, \quad (14)$$

$$\text{TGV}_{\alpha}^2(r^k) < M, \quad \|i - r^k - l^k\|_2^2 < M, \quad (15)$$

$$r^k \in \mathcal{B}_r, \quad l^k \in \mathcal{B}_l, \quad (16)$$

for $\forall k \in \mathbb{N}$. $\{l^k\}$ is uniformly bounded in $H^1(\Omega)$ by the bounded conditions (14). With the compact embedding of $H^1(\Omega)$ in $L^2(\Omega)$, it is easy to deduce that $\{l^k\}_{k \in \mathbb{N}}$ converges (up to a subsequence) in $H^1(\Omega)$ to l^* . As a consequence of the lower semicontinuity for $H^1(\Omega)$ norm, we obtain

$$\frac{c_1}{2} \|\nabla l^*\|_2^2 + \frac{c_3}{2} \|l^*\|_2^2 \leq \liminf_{k \rightarrow \infty} \left(\frac{c_1}{2} \|\nabla l^k\|_2^2 + \frac{c_3}{2} \|l^k\|_2^2 \right). \quad (17)$$

For $\{r^k\}$, we can get

$$\|r^k\|_2^2 = \|r^k + l^k - i + i - l^k\|_2^2 \leq \|r^k + l^k - i\|^2 + \|i - l^k\|_2^2 < M.$$

As $i \in L^\infty(\Omega)$ and $\{l^k\}$ is uniformly bounded in $L^2(\Omega)$, $\{r^k\}$ is also uniformly bounded in $L^2(\Omega)$. By the boundedness of Λ , $\{r^k\}$ is bounded in $L^1(\Omega)$ and bounded in $\text{BVG}_\alpha^2(\Omega)$ by the boundedness of $\text{TGV}_\alpha^2(r^k)$ in (15) and the lower semicontinuity. Similar with $\{l^k\}$, we can derive that $\{r^k\}_{k \in \mathbb{N}}$ converges (up to a subsequence) in $\text{BVG}_\alpha^2(\Omega)$ to $r^* \in \text{BVG}_\alpha^2(\Omega)$. By the lower semicontinuity of $\text{BVG}_\alpha^2(\Omega)$ norm, we obtain

$$\text{TGV}_\alpha^2(r^*) \leq \liminf_{k \rightarrow \infty} \text{TGV}_\alpha^2(r^k). \quad (18)$$

Since $l^k + r^k \rightarrow l^* + r^*$ in $L^2(\Omega)$ and recalling the lower semicontinuity for the $L^2(\Omega)$ norm, we also obtain

$$\frac{c_2}{2} \|i - r^* - l^*\|_2^2 \leq \liminf_{k \rightarrow \infty} \frac{c_2}{2} \|i - r^k - l^k\|_2^2. \quad (19)$$

From the above inequations (17), (18) and (19), we have

$$\min_{(r,l) \in \Lambda} E(r, l) = \liminf_{k \rightarrow \infty} E(r^k, l^k) \geq E(r^*, l^*).$$

Since the functional (11) is convex (by the convexity of TGV_α^2) and the constraint sets \mathcal{B}_r and \mathcal{B}_l are closed, the minima can be attained at (r^*, l^*) . Uniqueness is straightforward with the strict convexity of $E(r, l)$ with respect to (r, l) . This completes the proof. \square

4 Algorithm

In this section, we present the algorithm for solving the problem (12). Note that the two unknowns r and l are coupled in (12), and our idea is to use the alternating minimization scheme based on split Bregman iteration to fully decouple the subproblems for r and l , which yields a simple iteration procedure. The details of our algorithm are described in Algorithm 1.

4.1 Optimizing r of (P1)

To solve the minimization subproblem containing a TGV term in (P1), we adopt the fast split Bregman method [11, 12, 18, 40]. More precisely, we introduce auxiliary variables $(\mathbf{w}, \mathbf{v}; \mathbf{b}, \mathbf{d})$ and penalty parameters (θ_0, θ_1) , transforming (P1) into the following form:

Algorithm 1 Alternating minimization scheme

1. Initialization: $r^0 = 0, l^0 = i, k = 0$ and parameters c_1, c_2, c_3 .

2. For the k th iteration:

$$(P1) \quad r^{k+1} = \arg \min_{r \in \mathcal{B}_r} \left\{ \text{TGV}_\alpha^2(r) + \frac{c_2}{2} \int_\Omega (i - r - l^k)^2 dx \right\},$$

$$(P2) \quad l^{k+1} = \arg \min_{l \in \mathcal{B}_l} \int_\Omega \left\{ \frac{c_1}{2} |\nabla l|^2 + \frac{c_2}{2} (i - l - r^{k+1})^2 + \frac{c_3}{2} l^2 \right\} dx.$$

3. Go back to step 2 until $\epsilon_r = \frac{\|r^{k+1} - r^k\|_2}{\|r^{k+1}\|_2} \leq \epsilon_1$ and $\epsilon_l = \frac{\|l^{k+1} - l^k\|_2}{\|l^{k+1}\|_2} \leq \epsilon_2$.

$$\begin{aligned} \max_{\mathbf{b}, \mathbf{d}} \min_{\substack{r \in \mathcal{B}_r \\ \mathbf{p}, \mathbf{w}, \mathbf{v}}} \left\{ E_1(r, \mathbf{p}, \mathbf{w}, \mathbf{v}; \mathbf{b}, \mathbf{d}) = \frac{c_2}{2} \int_\Omega (i - r - l^k)^2 dx \right. \\ \left. + \int_\Omega \left(\alpha_1 |\mathbf{w}| + \frac{\theta_1}{2} (\mathbf{w} - \nabla r + \mathbf{p} - \mathbf{b})^2 \right) dx \right. \\ \left. + \int_\Omega \left(\alpha_0 |\mathbf{v}| + \frac{\theta_0}{2} (\mathbf{v} - \mathcal{E}(\mathbf{p}) - \mathbf{d})^2 \right) dx \right\}. \end{aligned} \quad (20)$$

We discrete the energy (20) on each pixel (details can be found in ‘‘Appendix’’) and break down (P1) into five steps. In the following, we will give a brief statement on solving each step.

Update of r^{t+1} To recover r^{t+1} , we take the partial derivative of the energy (20) with respect to r , which leads to the following Euler equation:

$$c_2 r - \theta_1 \Delta r = c_2 (i - l^k) - \theta_1 \operatorname{div}(\mathbf{w}^t + \mathbf{p}^t - \mathbf{b}^t), \quad (21)$$

The solution of this linear PDE is given by discrete Fourier transform \mathcal{F} , which namely

$$r^{t+\frac{1}{2}} = \Re \left(\mathcal{F}^{-1} \left(\frac{\mathcal{F}(G_r)}{\xi_r} \right) \right), \quad (22)$$

where $\mathcal{F}(G_r)$ and ξ_r are derived in ‘‘Appendix’’. Then, we apply the projection method to $r^{t+\frac{1}{2}}$, thus $r^{t+1} = \mathcal{P}_{\mathcal{B}_r}(r^{t+\frac{1}{2}})$.

Update of \mathbf{p}^{t+1} Similar with r^{t+1} , we derive the Euler equation with respect to $\mathbf{p}^t = (p_1^t, p_2^t)$. For each p_1^t and p_2^t , there are two discrete formulations as follows:

$$\left(\theta_1 - \theta_0 \partial_x^+ \partial_x^- - \frac{\theta_0}{2} \partial_y^+ \partial_y^- \right) p_1^{t+1} - \frac{\theta_0}{2} \partial y^+ \partial x^- p_2^{t+1} = h_1, \quad (23)$$

$$\left(\theta_1 - \frac{\theta_0}{2} \partial_x^+ \partial_x^- - \theta_0 \partial_y^+ \partial_y^- \right) p_2^{t+1} - \frac{\theta_0}{2} \partial y^+ \partial x^- p_1^{t+1} = h_2. \quad (24)$$

where the differential operators in (23) and (24) can be found in “Appendix” and

$$\begin{aligned} h_1 &= \theta_1(\partial_x^+ r^{t+1} + b_1^t - w_1^t) - \theta_0 \partial_x^+(v_{11}^t - d_{11}^t) \\ &\quad - \theta_0 \partial_y^+(v_3^t - d_3^t), \\ h_2 &= \theta_1(\partial_y^+ r^{t+1} + b_2^t - w_2^t) - \theta_0 \partial_y^+(v_{22}^t - d_{22}^t) \\ &\quad - \theta_0 \partial_x^+(v_3^t - d_3^t), \end{aligned}$$

By applying the discrete Fourier transform \mathcal{F} to both sides of (23) and (24), we get the following linear system

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} \mathcal{F}(p_1^{t+1}) \\ \mathcal{F}(p_2^{t+1}) \end{pmatrix} = \begin{pmatrix} \mathcal{F}(h_1) \\ \mathcal{F}(h_2) \end{pmatrix}, \quad (25)$$

and easily obtain the analytical solution of p_1^{t+1} and p_2^{t+1} with discrete inverse Fourier transform \mathcal{F}^{-1} .

Update of \mathbf{w}^{t+1} and \mathbf{v}^{t+1} The subproblems for \mathbf{w} and \mathbf{v} in (20) have closed form solutions using shrink operator as follows:

$$\mathbf{w}^{t+1} = \text{shrink}\left(\nabla r^{t+1} - \mathbf{p}^{t+1} + \mathbf{b}^t, \frac{c_1}{\theta_1}\right), \quad (26)$$

$$\mathbf{v}^{t+1} = \text{shrink}\left(\mathcal{E}(\mathbf{p}^{t+1}) + \mathbf{d}^t, \frac{c_2}{\theta_0}\right), \quad (27)$$

where the shrink operator is defined as

$$\text{shrink}(s, t) = \frac{s}{|s|_1} \max(|s|_1 - t, 0).$$

Update of \mathbf{b} and \mathbf{d} To enforce $\mathbf{w} = \nabla r - \mathbf{p}$ and $\mathbf{v} = \mathcal{E}(\mathbf{p})$, both Bregman iterative parameters \mathbf{b}^{t+1} and \mathbf{d}^{t+1} are updated as follows:

$$\begin{aligned} \mathbf{b}^{t+1} &= \mathbf{b}^t + \nabla r^{t+1} - \mathbf{p}^{t+1} - \mathbf{w}^{t+1}, \\ \mathbf{d}^{t+1} &= \mathbf{d}^t + \mathcal{E}(\mathbf{p}) - \mathbf{v}. \end{aligned} \quad (28)$$

Finally, the split Bregman algorithm for r subproblem (20) is summarized in Algorithm 2.

Algorithm 2 The split Bregman algorithm for r subproblem of (P1)

Initialization: Set $t = 0$, $(\mathbf{p}, \mathbf{w}, \mathbf{v}; \mathbf{b}, \mathbf{d}) = \mathbf{0}$ and parameters $(\alpha_0, \alpha_1, \theta_1, \theta_2) > 0$.

Repeat

- 1) Compute r^{t+1} according to (22),
- 2) Compute \mathbf{p}^{t+1} according to (25),
- 3) Compute \mathbf{w}^{t+1} according to (26),
- 4) Compute \mathbf{v}^{t+1} according to (27),
- 5) Update Bregman iterative parameter \mathbf{b} and \mathbf{d} according to (28),

Until $\frac{\|r^{t+1} - r^t\|_2}{\|r^{t+1}\|_2} < \epsilon_3$,

Return r^{t+1} .

4.2 Optimizing l of (P2)

The subproblem of (P2) is rather simple since it is a quadratic optimization under the closed constraint \mathcal{B}_l . First, Euler equation is derived with respect to l as follows:

$$-c_1 \Delta l + (c_2 + c_3)l = c_2(i - r^{k+1}), \quad (29)$$

then solved efficiently by discrete Fourier transform \mathcal{F} under the periodic boundary condition, which namely

$$l^{k+\frac{1}{2}} = \Re\left(\mathcal{F}^{-1}\left(\frac{\mathcal{F}(G_l)}{\xi_l}\right)\right),$$

where $\mathcal{F}(G_l)$ and ξ_l are derived in “Appendix”. Finally, we apply the projection method to $l^{k+\frac{1}{2}}$, thus $l^{k+1} = \mathcal{P}_{\mathcal{B}_l}(l^{k+1})$.

The final estimated reflectance and illumination are obtained by $R = \exp(r)$ and $L = \exp(l)$, respectively. In the next section, the experimental results are presented to show the effectiveness of the proposed algorithm.

5 Experimental results and comparison

In this section, we compare the performance of our proposed model with several variational Retinex methods including VF [28], L1 [42], TVM [51], HoTVL1 [38] and WVM [14], the state-of-the-art HDR tone mapping method [19] and supervised deep learning-based method [70] and consider the decomposition problem on several classic test images including synthetic examples, natural images with low contrast or non-uniform lighting and medical images. All experiments are executed with MATLAB R2011b on a windows 7 platform with an Intel Corei7 at 3.2 GHz and 8 GB RAM.

For all tests, the recovered r of the VF method is obtained via $i - l$ since the reflection function of VF is not considered, and the recovered l of the L1 method is obtained via $i - r$. For fair comparison, the parameters are set to be optimal according to their papers. In our algorithm, we fixed $c_3 = 1e^{-8}$, $\alpha_0 = 1$, $\alpha_1 = 0.5$, $\theta_0 = \theta_1 = 5$ and the stopping parameters $\epsilon_1 = \epsilon_2 = 5e^{-4}$. We use the HSV Retinex model for color images with constraint sets $\mathcal{B}_r = \mathcal{B}_l = [\log(1e^{-10}), 0]$ and omit the logarithm and exponent process for gray images just as suggested in [50].

5.1 Synthetic images

We start with two different synthetic examples, shown in Figs. 3 and 4, respectively. In Fig. 3, we get the darker color image by changing luminance of the ground truth. In Fig. 4, we make a piecewise constant gray image r and a piecewise smooth gray image l , then get the simulated image i by summation of them, namely $i = r + l$.

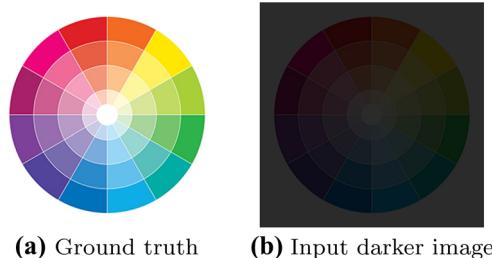


Fig. 3 Synthetic example 1. **a** Ground truth. **b** Input darker image

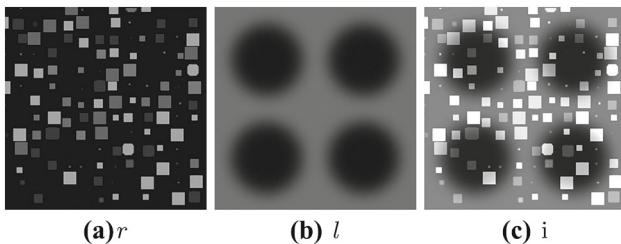


Fig. 4 Synthetic example 2. **a** Synthetic reflectance image r ; **b** synthetic illumination/shadow image l and **c** composed image $i = r + l$

The decomposition results are shown in Figs. 5 and 6, respectively. To measure the accuracy of the estimated r , the S-CIELAB color metric [71] is adopted by drawing spatial locations and histogram distribution of errors. The first

and second row in Fig. 5 shows the estimated reflectance and illumination results of different methods, respectively. The third row in Fig. 5 marks the pixels whose S-CIELAB errors between the ground truth and estimated reflectance of different methods are > 10 units, and the last row shows the histogram distribution of S-CIELAB errors between the ground truth and the estimated reflectance of different methods. As can be seen from the spatial locations of the errors, the green area of our result is smaller than those of other five methods. From the histogram distribution of errors, it indicates our result achieves the lowest maximum and average S-CIELAB error in agreement with our empirical assessment. This example demonstrates that our proposed model can recover more accurate result in dealing with darker image. From Fig. 6, we can observe that TGVH1, L1 and HoTVL1 produce the better visual results than other methods, where the estimated r and l better approximate the ground truth in some properties. Intuitively, VF, WVM and TVM either take the illumination information as details of reflectance or lose many sharp features of reflectance, such as the edges of rectangles shown in Fig. 6.

For these two examples, we both set $c_1 = c_2 = 20$, and the constraint sets for the second example are $\mathcal{B}_r = \mathcal{B}_l = [0, 255]$.

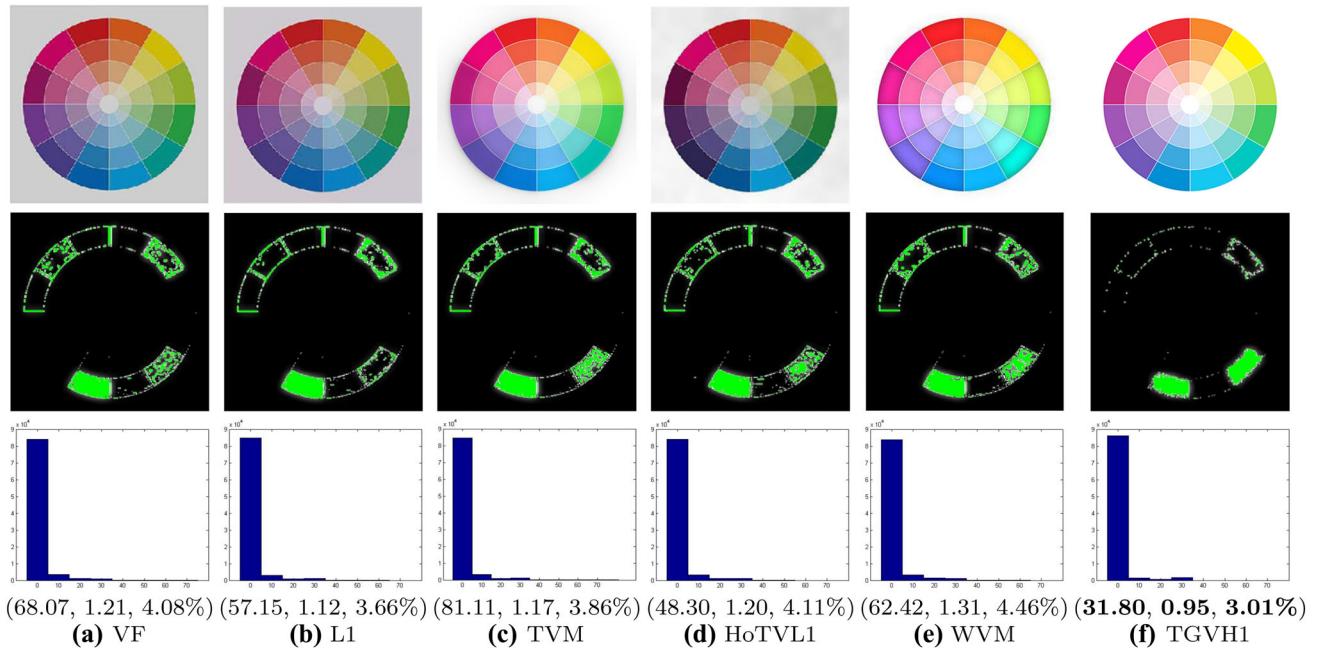


Fig. 5 Decomposition comparison of synthetic example 1. First row: estimated reflectance images by VF, L1, TVM, HoTVL1, WVM and proposed method, respectively. Second row: the spatial distribution of S-CIELAB errors between the ground truth and estimated reflectance components by VF, L1, TVM, HoTVL1, WVM and proposed method, respectively. Numbers in bracket denote the maximum, average of the S-CIELAB and the proportion of pixels whose S-CIELAB error are > 10 units, respectively

respectively. Third row: the histogram distribution of S-CIELAB errors between the ground truth and estimated reflectance components by VF, L1, TVM, HoTVL1, WVM and proposed method, respectively. Numbers in bracket denote the maximum, average of the S-CIELAB and the proportion of pixels whose S-CIELAB error are > 10 units, respectively

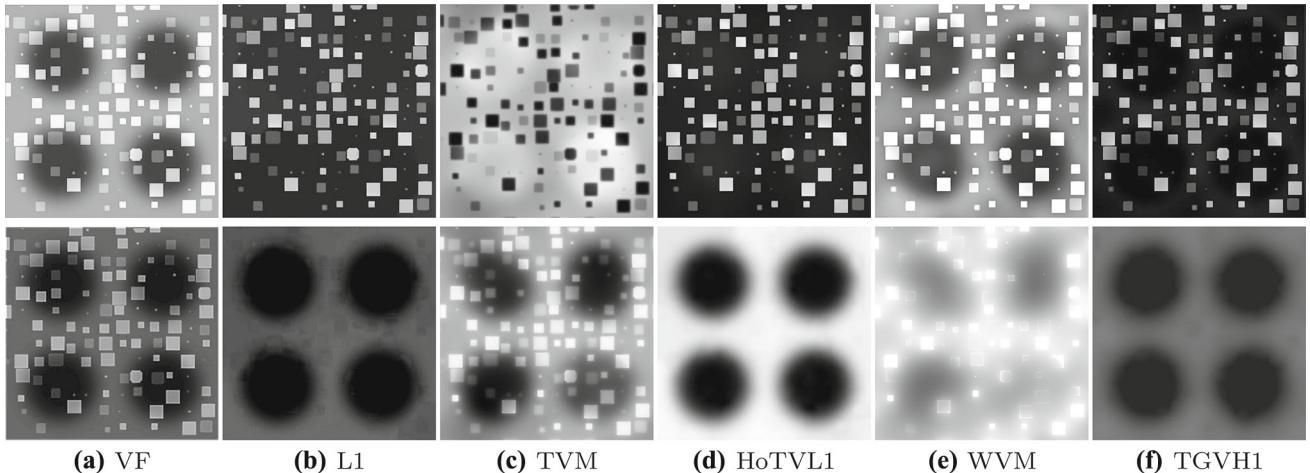


Fig. 6 Decomposition comparison of synthetic example 2. First row: estimated reflectance images by VF, L1, TVM, HoTVL1, WVM and proposed method, respectively. Second row: estimated illumination images by VF, L1, TVM, HoTVL1, WVM and proposed method, respectively

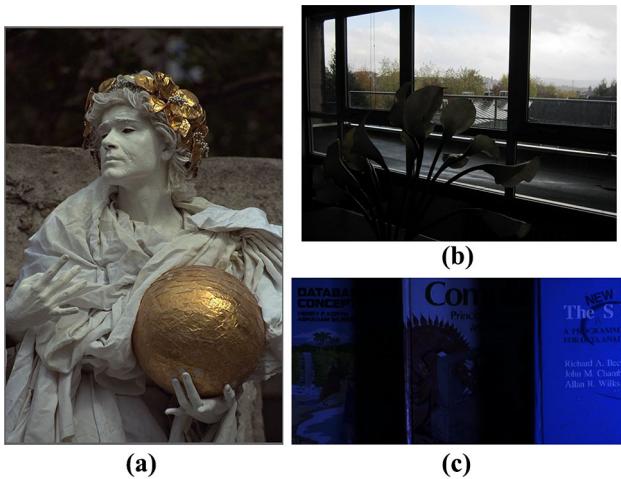


Fig. 7 Natural images

5.2 Natural image color correction

To demonstrate the performance of our proposed model on natural image color correction, three tests are done on natural images in Fig. 7, and the comparison results of different methods are presented in Figs. 8, 9 and 10.

In these three figures, we show the comparisons of different methods both on reflectance and illumination. One can observe that details of the high-contrast parts are more improved by using our proposed algorithm than other methods. The improved details include the wrinkle on the man's face in Fig. 8, the tree leaves and the radiators behind flower leaves in Fig. 9, and the plants and characters in Fig. 10. The corresponding zoomed parts are shown with enlargements in red rectangles. Note that the TVM method always gives overenhanced reflectance with less details and adds the details to recovered illumination. L1 and HoTVL1 suf-

fer from the loss of reflectance details and contrast due to the penalization on the magnitude of image gradient. As VF does not consider reflectance directly, so the $i - l$ always performs poorly, especially for low-contrast color images. WVM can estimate reflectance with finer details due to the modification of the weight in log-transform domain, but it is failed for some low-contrast color images, shown in Fig. 10. However, TGVH1 can successfully reveal reflectance with many details in shadows and illumination with smooth property, increasing the contrast simultaneously.

In addition, since the ground truth of the reflectance image is unknown, we utilize the natural image quality evaluator (NIQE) [48] to evaluate the results of different methods. A lower NIQE value represents a higher image quality and the best results are boldfaced, as shown in Table 1. It demonstrates that our method achieves a lower average than other methods, consistent with our empirical assessment that our method outputs the highest quality of reflectance images.

For the first two examples, we both set $c_1 = c_2 = 50$, and $c_1 = c_2 = 1$ for the third example.

5.3 Medical image bias field correction

Medical images such as MRI are always corrupted by bias fields due to non-uniform illuminations, so the correction of the bias field by removing the light effect caused by illumination is important for clinical diagnosis. Here, we test three brain MRI images in Fig. 11 and apply different models to recover a detail-preserved reflectance shown in Fig. 12. For fair comparison, we replace L1 by the smoothed L1 method [42] to recover many subtle details followed [42]. In the original MRI images (see in Fig. 11), there is apparent bias field effect at the neck area we hardly observe. In Fig. 12, from left to right are the results of six methods where tissues

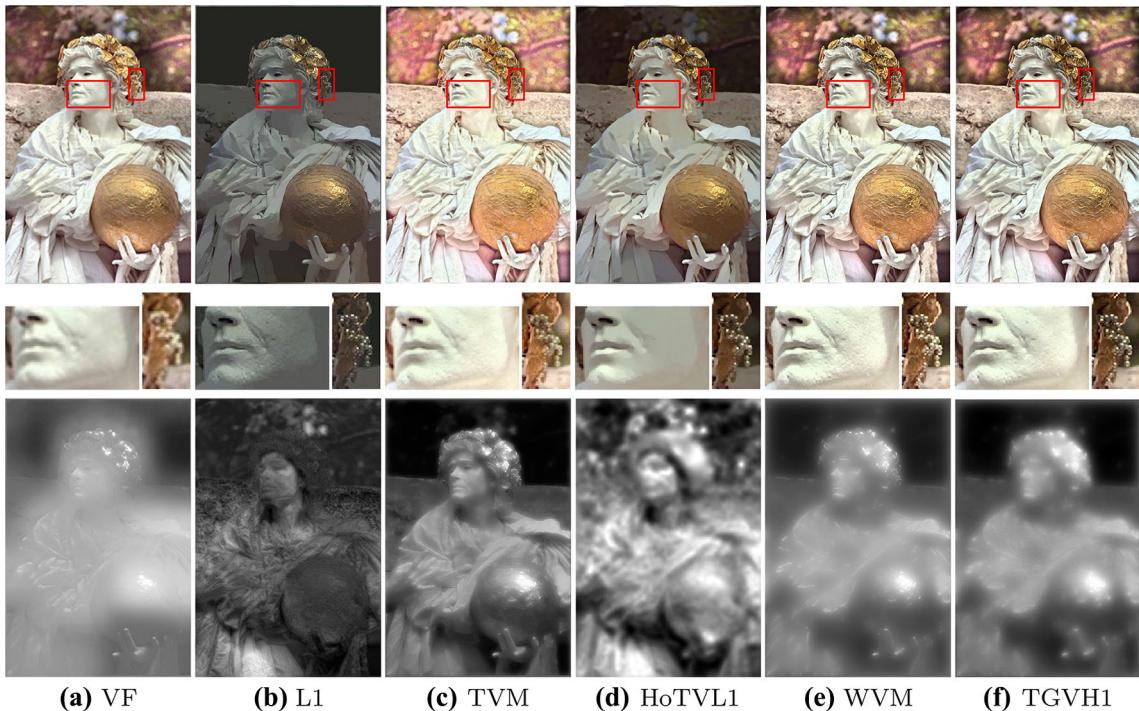


Fig. 8 Decomposition comparison of natural example 1. First row: estimated reflectance images by VF, L1, TVM, HoTVL1, WVM and proposed method, respectively. Second row: estimated illumination images by VF, L1, TVM, HoTVL1, WVM and proposed method, respectively

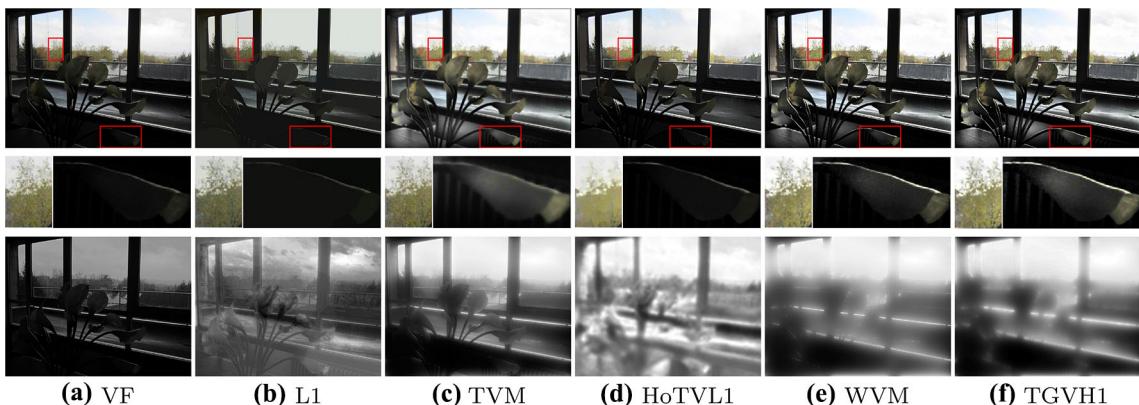


Fig. 9 Decomposition comparison of natural example 2. First row: estimated reflectance images by VF, L1, TVM, HoTVL1, WVM and proposed method, respectively. Second row: estimated illumination images by VF, L1, TVM, HoTVL1, WVM and proposed method, respectively

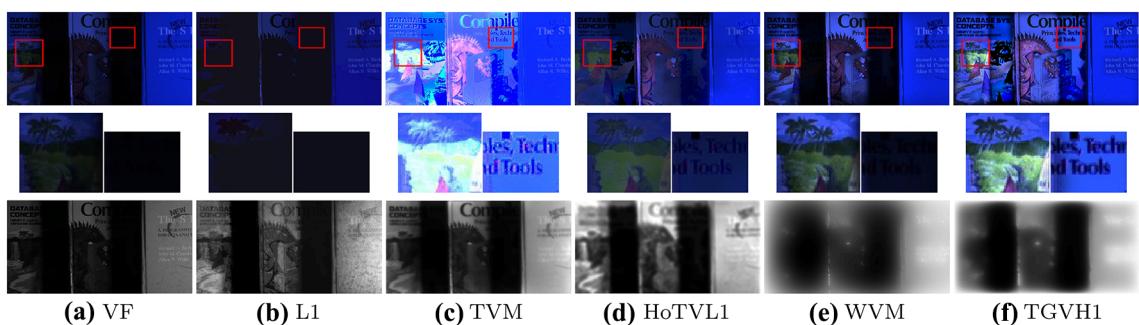
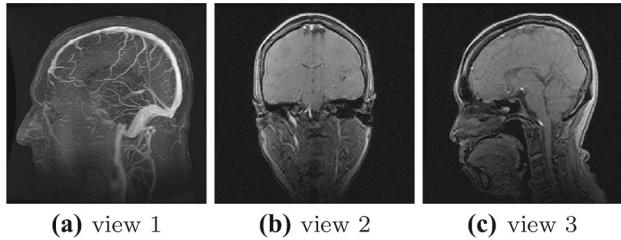


Fig. 10 Decomposition comparison of natural example 3. First row: estimated reflectance images by VF, L1, TVM, HoTVL1, WVM and proposed method, respectively. Second row: estimated illumination images by VF, L1, TVM, HoTVL1, WVM and proposed method, respectively

Table 1 NIQE values of Figs. 8, 9 and 10

Images	Figure 8	Figure 9	Figure 10	Average
Input	2.9320	4.4922	5.5415	4.3219
VF	3.5981	3.7245	5.3905	4.2377
L1	5.5415	7.0833	8.8603	7.1617
TVM	2.8592	5.3776	4.7948	7.0895
HoTVL1	3.7471	4.7656	4.7295	4.4141
WVM	2.7344	2.6968	4.1582	3.1965
TGVH1	2.7292	2.7005	4.0316	3.1538

**Fig. 11** Medical examples

and vessels are significantly improved. It is obvious that the HoTVL1 method gets the piecewise constant results where some small details are lost, and the WVM method introduces some noises, especially in the black background. Other

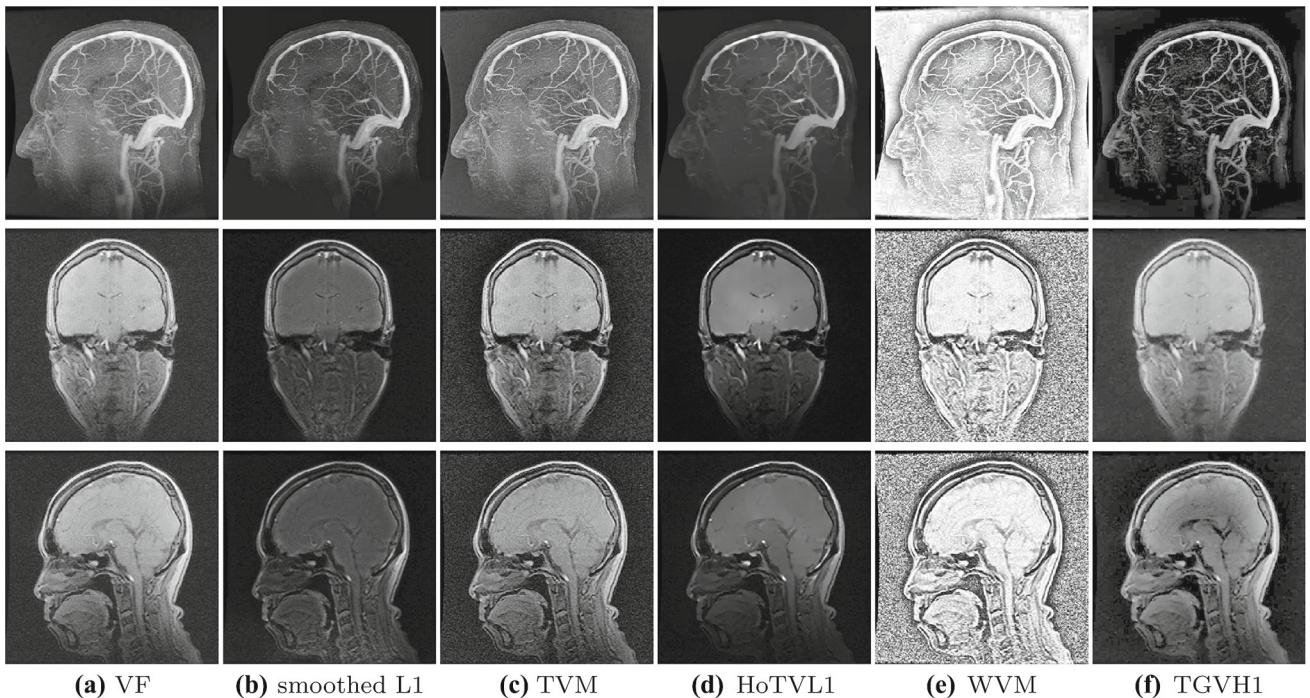
methods produce good bias field corrections, especially for the details near the lower neck area and the edge of the MRI slice are significantly visible by TGVH1 with the suppression of noises in the background.

For these two examples, we both set $c_1 = c_2 = 0.5$, and the constraint sets for these examples are all set to $\mathcal{B}_r = \mathcal{B}_l = [-255, 0]$.

5.4 Shadow removal of color images

The final explanatory example is the removal of shadowed color images. Figure 13 gives two color images with different kinds of shadow, and the comparison results of different methods are shown in Figs. 14 and 15. Note that the L1 method can eliminate shadow effectively, but always recovers reflectance with big color difference, mainly due to the value of one pixel fixed in solving a Poisson equation for the sake of a unique solution. VF, TVM and WVM can perform better for images with milder shadows, but fail to remove the adequate shadows shown in Fig. 15. The HoTVL1 and TGVH1 method can give visually accurate results where the shadow effect can be partially removed. Moreover, TGVH1 can simulate the shadows spread all over the whole image, including the shadow area of two signboards in Wall image.

For the first example, we set $c_1 = c_2 = 50$, and $c_1 = 0.1, c_2 = 5$ for the second example.

**Fig. 12** Comparison of Medical image bias field correction. First row: the results of view 1 by VF, L1, TVM, HoTVL1, WVM and proposed method, respectively. Second row: the results of view 2 by VF, L1,

TVM, HoTVL1, WVM and proposed method, respectively. Third row: the results of view 3 by VF, L1, TVM, HoTVL1, WVM and proposed method, respectively

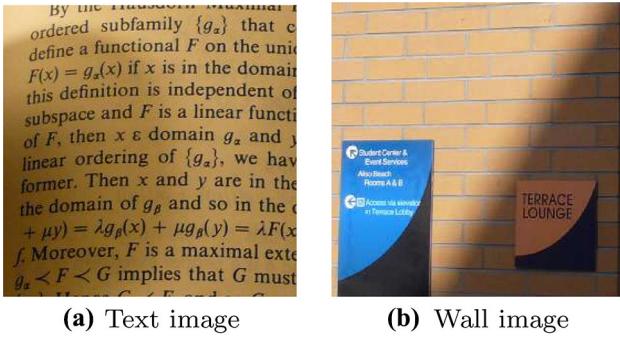


Fig. 13 Color images for shadow removal

5.5 Comparison with HDR tone mapping method

We compare our method with the state-of-the-art HDR tone mapping method JCAS [19] aimed to reproduce a low dynamic range (LDR) image with fine details and colors of the HDR counterpart. The source code is obtained from the original authors, and parameters are set following the paper.

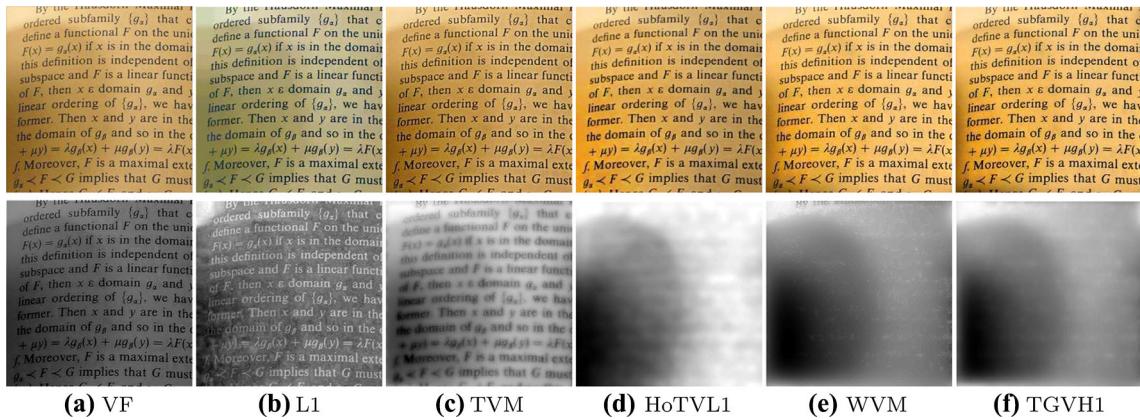


Fig. 14 Comparison of shadow removal on Text image. First row: estimated reflectance images by VF, L1, TVM, HoTVL1, WVM and proposed method, respectively. Second row: estimated illumination images by VF, L1, TVM, HoTVL1, WVM and proposed method, respectively

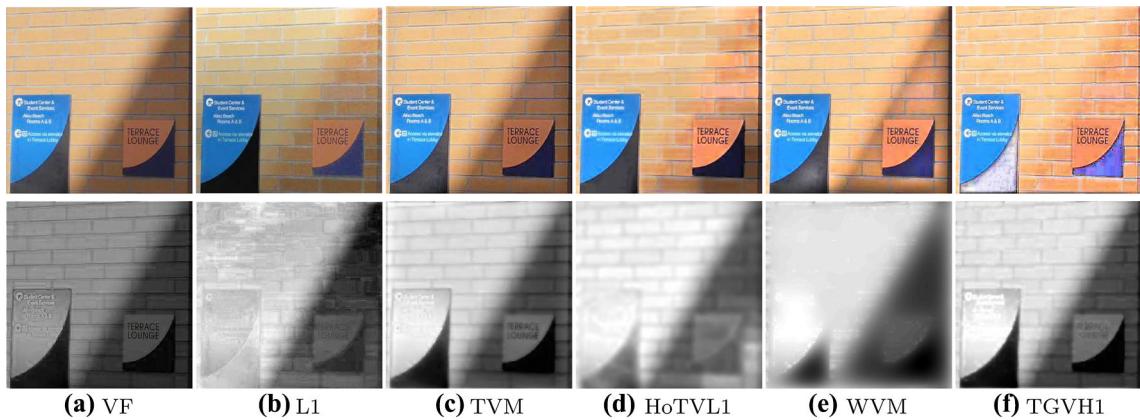
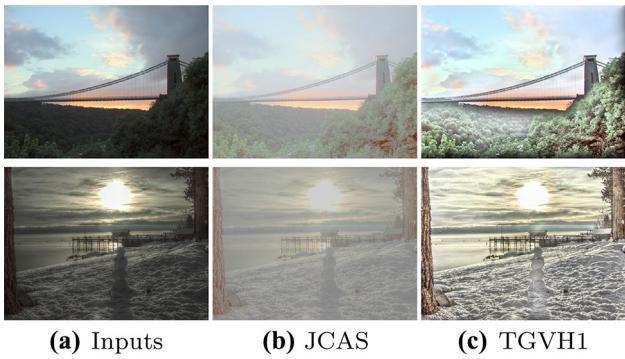
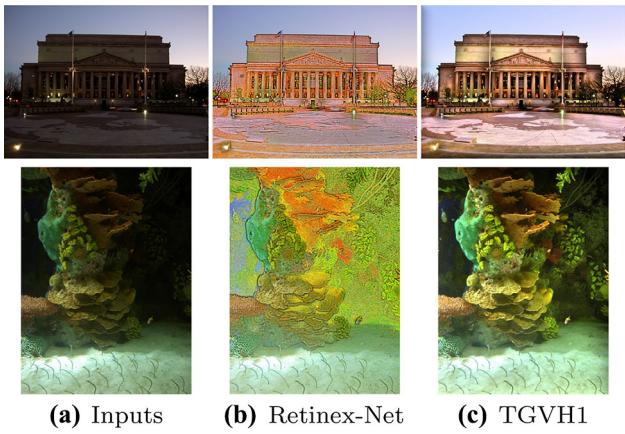


Fig. 15 Comparison of shadow removal on Wall image. First row: estimated reflectance images by VF, L1, TVM, HoTVL1, WVM and proposed method, respectively. Second row: estimated illumination images by VF, L1, TVM, HoTVL1, WVM and proposed method, respectively

Figure 16 shows visual comparison on two natural images. It is easy to see that JCAS can recover most of the details in the HDR image, but the results are too overexposed and blurry, while our method generates higher quality reflectance with more details and faithful colors benefited from the edge-preserving property of TGV regularization.

5.6 Comparison with supervised deep learning-based method

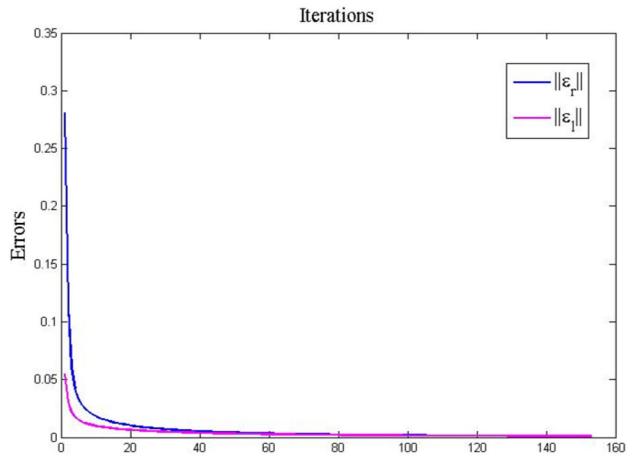
We also compare our proposed method with Retinex-Net [70], which is the most state-of-the-art supervised deep learning-based method for low-light image enhancement built on Retinex theory, on real-scene images from public DICM [35] dataset collected 69 images with commercial digital cameras. We get all Retinex-Net's results with the trained model provided by author. Figure 17 shows visual comparison on two natural images. As shown in each result, our method estimates the reflectance with more details without overexposure, especially for the objects buried in dark

**Fig. 16** Comparison with JCAS [19]**Fig. 17** Comparison with Retinex-Net [70]**Table 2** Average NIQE values on several public datasets

	DICM	Fusion	LIME	MEF	NPE
Input	3.8597	4.2050	4.3521	5.1891	3.7388
Retinex-Net	4.6753	4.3809	4.6609	5.7145	4.7079
TGVH1	2.7632	2.7814	3.0709	3.3918	2.9656

lightness, which benefits from the edge-preserving property of TGV regularization. In addition, our method is succeeded in suppressing noise in dark areas, while Retinex-Net contains so many noises. In other words, our proposed method outperforms Retinex-Net.

We also evaluate the reflectance component of our method quantitatively on real-scene images from public DICM [35], Fusion [65], LIME [20], MEF [41] and NPE [66]. Table 2 shows the corresponding average NIQE values of the five datasets and demonstrates that our method has the lower NIQE value on natural images than Retinex-Net, which indicates that our model estimates the reflectance with high quality.

**Fig. 18** Change of ϵ_r and ϵ_l with the number of iterations on a 512×512 image

5.7 Convergence

Since our energy function is locally convex with respect to r and l , it is effective to apply alternating minimization scheme to decrease the energy monotonically. However, it should be mentioned that it is difficult to prove the convergence of Algorithm 1 because of the bound box constraints on r and l . Furthermore, in real computations, we do not employ many inner loop iterations of Algorithm 2 and in the numerical tests, we set the max iterations as 5. Figure 18 shows the change of ϵ_r , ϵ_l with the number of iterations on a 512×512 image. From the error curve, it is clear that both errors drop severely during the optimization and asymptotically goes toward 0 with increasing number of iterations.

5.8 Parameters impact

We do some tests about the parameters impact on the proposed model (Eq. (12)). c_1 and c_2 are related to the regularization of illumination, and we found the algorithm is relatively stable when they are equal for most examples in our tests and the example is done with $c_1 = c_2 = 50$ shown in Fig. 8. To test the impact of different parameters, respectively, we set both parameters to 1 and 5 to compare. As can be seen in Fig. 19, details of estimated reflectance are fuzzed since the penalty of illumination decreases as c_1 and c_2 decreases, leading to the estimated illumination with more details while reflectance with poor details. Conversely, we can get a detail-preserved reflectance and smooth illumination.

5.9 Computational time

As an important factor of an algorithm, computational time of different methods is compared. Among the methods, the VF method is the fastest and the L1 method is the sec-

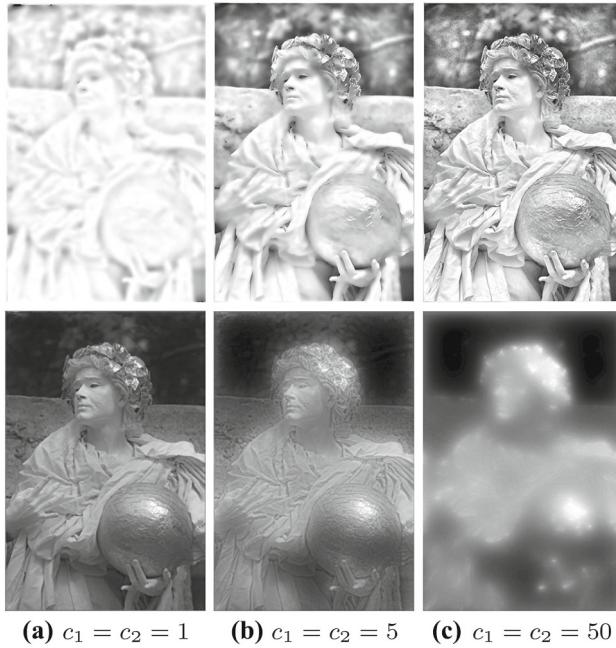


Fig. 19 Examples of parameters impact. First row: estimated reflectance images. Second row: estimated illumination images

ond since they both consider only one variable. The TVM and WVM methods are the third and fourth where both r and l are considered. HoTVL1 and our method both are time-consuming mainly because the former involves the second-order TV regularization and the latter involves TGV regularization. Specifically, our algorithm is a two-step alternating minimization scheme including r and l update, and the most time-consuming part is the r update which consists of three subproblems. For example, in Fig. 10 with image size 713×357 , it takes 7.6 s for the VF method, 9.5 s for the L1 method with 112 iterations, 71.3 s for the TVM method with 175 iterations, 466.5 s for the WVM method with 249 iterations, 920.1 s for the HoTVL1 method with 5342 iterations and 890.2 s for our method with 477 iterations. Since the experiments are performed by unoptimized MATLAB implementation, the computational time can be improved by C programming or trying other algorithms.

6 Conclusion

In this paper, we have presented a reflectance and illumination decomposition model for the Retinex problem via total generalized variation regularization and H^1 decomposition, which nicely separates the smoother global illumination from richly detailed reflectance. We proved the existence and uniqueness of the minimizers for the proposed model and provided an alternating minimization algorithm based on split Bregman iteration to handle the convex optimiza-

tion under some constraints. Compared with state-of-the-art methods, the proposed algorithm performed effectively to capture more detailed structures or features of the reflectance and extract HVS preferable illumination for both inhomogeneous background removal and color shadow correction.

In the future, we hope to consider the decomposition problem on some low-quality images, such as severe noised, blurred images in low-light conditions. How to extend these techniques to the digital images of wall paintings which are captured in the dark conditions will also be a future work.

Acknowledgements We would like to thank Michael K. Ng [51], Jingwei Liang [38], Xueyang Fu [14] and Chen Wei [70] for sharing their code and the reviewers of this manuscript for their helpful comments and suggestions. This work is supported by the National Natural Science Foundation of China (61802279, 61602341) and the National Natural Science Foundation of Tianjin (18JCQNJC00100, 17JCQNJC00600).

Appendix

In this appendix, we introduce detailed discretization of first-order and second-order differential operators using finite difference scheme.

The first-order forward and backward difference schemes are first given. Let $\Omega \rightarrow \mathbb{R}^{M \times N}$ denote the two-dimensional grayscale image space with size M and N . The coordinates x and y are oriented along columns and rows, respectively. So the first-order forward differences of u at point (i, j) along x and y directions are, respectively,

$$\partial_x^+ u_{i,j} = \begin{cases} u_{i,j+1} - u_{i,j} & \text{if } 1 \leq i \leq M, 1 \leq j < N, \\ u_{i,1} - u_{i,j} & \text{if } 1 \leq i \leq M, j = N. \end{cases}$$

$$\partial_y^+ u_{i,j} = \begin{cases} u_{i+1,j} - u_{i,j} & \text{if } 1 \leq i < M, 1 \leq j \leq N, \\ u_{1,j} - u_{i,j} & \text{if } i = M, 1 \leq j \leq N. \end{cases}$$

The first-order backward differences are, respectively,

$$\partial_x^- u_{i,j} = \begin{cases} u_{i,j} - u_{i,j-1} & \text{if } 1 \leq i \leq M, 1 < j \leq N, \\ u_{i,j} - u_{i,N} & \text{if } 1 \leq i \leq M, j = 1. \end{cases}$$

$$\partial_y^- u_{i,j} = \begin{cases} u_{i,j} - u_{i-1,j} & \text{if } 1 < i \leq M, 1 \leq j \leq N, \\ u_{i,j} - u_{M,j} & \text{if } i = 1, 1 \leq j \leq N. \end{cases}$$

And the discrete second-order derivatives $\partial_x^- \partial_x^+ u$, $\partial_x^+ \partial_x^- u$, $\partial_y^- \partial_y^+ u$ and $\partial_y^+ \partial_y^- u$ at point (i, j) can be written by the corresponding compositions of the discrete first-order derivative, as follows

$$\begin{aligned}\partial_x^+ \partial_x^- u_{i,j} &= \begin{cases} u_{i,N} - 2u_{i,j} + u_{i,j+1} & \text{if } 1 \leq i \leq M, j = 1, \\ u_{i,j-1} - 2u_{i,j} + u_{i,j+1} & \text{if } 1 \leq i \leq M, 1 < j < N, \\ u_{i,j-1} - 2u_{i,j} + u_{i,1} & \text{if } 1 \leq i \leq M, j = N. \end{cases} \\ &= \partial_x^- \partial_x^+ u_{i,j} \\ \partial_y^+ \partial_y^- u_{i,j} &= \begin{cases} u_{M,j} - 2u_{i,j} + u_{i+1,j} & \text{if } i = 1, 1 \leq j \leq N, \\ u_{i-1,j} - 2u_{i,j} + u_{i+1,j} & \text{if } 1 < i < M, 1 \leq j \leq N, \\ u_{i-1,j} - 2u_{i,j} + u_{1,j} & \text{if } i = M, 1 \leq j \leq N. \end{cases} \\ &= \partial_y^- \partial_y^+ u_{i,j}\end{aligned}$$

Thus, the gradient, symmetrized derivative, divergence and Laplace can be discretized as follows, respectively:

$$\nabla u = (\partial_x^+ u, \partial_y^+ u), \quad (30)$$

$$\mathcal{E}(\mathbf{p}) = \begin{pmatrix} \partial_x^- p_1 & \frac{\partial_y^- p_1 + \partial_x^- p_2}{2} \\ \frac{\partial_y^- p_1 + \partial_x^- p_2}{2} & \partial_y^- p_2, \end{pmatrix} \quad (31)$$

$$\operatorname{div}(\mathbf{p}) = \partial_x^- p_1 + \partial_y^- p_2, \quad (32)$$

$$\Delta u_{i,j} = \operatorname{div}(\partial u_{i,j}) = \partial_x^- \partial_x^+ u_{i,j} + \partial_y^- \partial_y^+ u_{i,j}. \quad (33)$$

So the $\mathcal{F}(G_r)$ and ξ_r in Eq. (22) can be written as

$$\begin{aligned}\mathcal{F}(G_r) &= \beta \mathcal{F}(i - l^k) - \theta_1 \mathcal{F}\{\partial_x^-(w_{1i,j} - b_{1i,j}) \\ &\quad + \partial_y^-(w_{2i,j} - b_{2i,j})\}.\end{aligned} \quad (34)$$

and

$$\begin{aligned}\xi_r &= \beta - \theta_1 \mathcal{F}(\partial_x^- \partial_x^+ + \partial_y^- \partial_y^+), \\ &= \beta - 2\theta_1 \left(\cos \frac{2\pi n}{N} + \cos \frac{2\pi m}{M} - 2 \right).\end{aligned} \quad (35)$$

References

1. Blake, A.: Boundary conditions for lightness computation in mondrian world. *Comput. Vis. Graph. Image Process.* **32**(3), 314–327 (1985)
2. Bredies, K., Kunisch, K., Pock, T.: Total generalized variation. *SIAM J. Imaging Sci.* **3**(3), 492–526 (2010)
3. Bredies, K., Kunisch, K., Valkonen, T.: Properties of 11-tgv2: the one-dimensional case. *J. Math. Anal. Appl.* **398**(1), 438–454 (2013)
4. Chambolle, A., Lions, P.L.: Image recovery via total variation minimization and related problems. *Numer. Math.* **76**(2), 167–188 (1997)
5. Chang, H., Ng, M.K., Wang, W., Zeng, T.: Retinex image enhancement via a learned dictionary. *Opt. Eng.* **54**(1), 013107–013107 (2015)
6. Cheng, M.H., Huang, T.Z., Zhao, X.L., Ma, T.H., Huang, J.: A variational model with hybrid hyper-Laplacian priors for retinex. *Appl. Math. Model.* **66**, 305–321 (2019)
7. Choi, D.H., Jang, I.H., Kim, M.H., Kim, N.C.: Color image enhancement based on single-scale retinex with a jnd-based nonlinear filter. In: Proceedings of 2007 IEEE International Symposium on Circuits and Systems, pp. 3948–3951. IEEE (2007)
8. Choi, D.H., Jang, I.H., Kim, M.H., Kim, N.C.: Color image enhancement using single-scale retinex based on an improved image formation model. In: Proceedings of the 16th European Signal Processing Conference, pp. 1–5. IEEE (2008)
9. Ciurea, F., Funt, B.: Tuning retinex parameters. *J. Electron. Imaging* **13**(1), 48–57 (2004)
10. Cooper, T.J., Baqai, F.A.: Analysis and extensions of the Frankle–McCann retinex algorithm. *J. Electron. Imaging* **13**(1), 85–93 (2004)
11. Duan, J., Pan, Z., Yin, X., Wei, W., Wang, G.: Some fast projection methods based on Chan–Vese model for image segmentation. *EURASIP J. Image Video Process.* **2014**(1), 7–7 (2014)
12. Duan, J., Pan, Z., Zhang, B., Liu, W., Tai, X.C.: Fast algorithm for color texture image inpainting using the non-local CTV model. *J. Global Optim.* **62**(4), 853–876 (2015)
13. Frankle, J.A., McCann, J.J.: Method and apparatus for lightness imaging (1983). US Patent 4384336
14. Fu, X., Zeng, D., Huang, Y., Zhang, X.P., Ding, X.: A weighted variational model for simultaneous reflectance and illumination estimation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2782–2790 (2016)
15. Funt, B., Ciurea, F., McCann, J.: Retinex in matlab. In: Proceedings of Color and Imaging Conference, vol. 2000, pp. 112–121. Society for Imaging Science and Technology (2000)
16. Funt, B.V., Drew, M.S., Brockington, M.: Recovering shading from color images. In: Proceedings of European Conference on Computer Vision, pp. 124–132. Springer (1992)
17. Gao, Y., Hu, H.M., Li, B., Guo, Q.: Naturalness preserved nonuniform illumination estimation for image enhancement based on retinex. *IEEE Trans. Multimed.* **20**(2), 335–344 (2017)
18. Goldstein, T., Osher, S.: The split bregman method for 11-regularized problems. *SIAM J. Imaging Sci.* **2**(2), 323–343 (2009)
19. Gu, S., Meng, D., Zuo, W., Zhang, L.: Joint convolutional analysis and synthesis sparse representation for single image layer separation. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 1708–1716 (2017)
20. Guo, X., Li, Y., Ling, H.: Lime: Low-light image enhancement via illumination map estimation. *IEEE Trans. Image Process.* **26**(2), 982–993 (2016)
21. Horn, B.K.: Determining lightness from an image. *Comput. Graph. Image Process.* **3**(4), 277–299 (1974)
22. Hou, G., Wang, G., Pan, Z., Huang, B., Yang, H., Yu, T.: Image enhancement and restoration: state of the art of variational retinex models. *IAENG Int. J. Comput. Sci.* **44**(4), 445–455 (2017)
23. Jiang, B., Woodell, G.A., Jobson, D.J.: Novel multi-scale retinex with color restoration on graphics processing unit. *J. Real-Time Image Proc.* **10**(2), 239–253 (2015)
24. Jobson, D.J., Rahman, Z., Woodell, G.A.: A multiscale retinex for bridging the gap between color images and the human observation of scenes. *IEEE Trans. Image Process.* **6**(7), 965–976 (1997)
25. Jobson, D.J., Rahman, Z., Woodell, G.A.: Properties and performance of a center/surround retinex. *IEEE Trans. Image Process.* **6**(3), 451–462 (1997)
26. Juha, K.: Sobolev spaces. Aalto University (2017)
27. Kang, M., Kang, M., Jung, M.: Total generalized variation based denoising models for ultrasound images. *J. Sci. Comput.* **72**(1), 172–197 (2017)
28. Kimmel, R., Elad, M., Shaked, D., Keshet, R., Sobel, I.: A variational framework for retinex. *Int. J. Comput. Vis.* **52**(1), 7–23 (2003)

29. Knoll, F., Bredies, K., Pock, T., Stollberger, R.: Second order total generalized variation (TGV) for MRI. *Magn. Reson. Med.* **65**(2), 480–491 (2011)
30. Land, E.H.: The retinex theory of color vision. *Sci. Am.* **237**(6), 108–129 (1977)
31. Land, E.H.: Recent advances in retinex theory and some implications for cortical computations: color vision and the natural image, pp. 5163–5169. National Academy of Sciences (1983)
32. Land, E.H.: Recent advances in retinex theory. In: *Central and Peripheral Mechanisms of Colour Vision*, pp. 5–17. Springer (1985)
33. Land, E.H.: An alternative technique for the computation of the designator in the retinex theory of color vision, pp. 3078–3080. National Academy Sciences (1986)
34. Land, E.H., McCann, J.J.: Lightness and retinex theory. *JOSA A* **61**(1), 1–11 (1971)
35. Lee, C., Lee, C., Kim, C.S.: Contrast enhancement based on layered difference representation. In: *2012 IEEE International Conference on Image Processing*, pp. 965–968. IEEE (2012)
36. Lei, L., Zhou, Y., Li, J.: An investigation of retinex algorithms for image enhancement. *J. Electron.* **24**(5), 696–700 (2007)
37. Li, H., Zhang, L., Shen, H.: A perceptually inspired variational method for the uneven intensity correction of remote sensing images. *IEEE Trans. Geosci. Remote Sens.* **50**(8), 3053–3065 (2012)
38. Liang, J., Zhang, X.: Retinex by higher order total variation l^1 decomposition. *J. Math. Imaging Vis.* **52**(3), 345–355 (2015)
39. Lore, K.G., Akintayo, A., Sarkar, S.: Llnet: a deep autoencoder approach to natural low-light image enhancement. *Pattern Recogn.* **61**, 650–662 (2017)
40. Lu, W., Duan, J., Qiu, Z., Pan, Z., Liu, R.W., Bai, L.: Implementation of high-order variational models made easy for image processing. *Math. Methods Appl. Sci.* **39**(14), 4208–4233 (2016)
41. Ma, K., Zeng, K., Wang, Z.: Perceptual quality assessment for multi-exposure image fusion. *IEEE Trans. Image Process.* **24**(11), 3345–3356 (2015)
42. Ma, W., Morel, J.M., Osher, S., Chien, A.: An l_1 -based variational model for retinex theory and its application to medical images. In: *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 153–160. IEEE (2011)
43. Ma, W., Osher, S.: A TV bregman iterative model of retinex theory. *Inverse Probl. Imaging* **6**(4), 697–708 (2012)
44. Ma, Y., Feng, X., Jiang, X., Xia, Z., Peng, J.: Intrinsic image decomposition: A comprehensive review. In: *International Conference on Image and Graphics*, pp. 626–638. Springer (2017)
45. Marini, D., Rizzi, A.: A computational approach to color adaptation effects. *Image Vis. Comput.* **18**(13), 1005–1014 (2000)
46. Maz'ya, V.: Sobolev Spaces. Springer, Berlin (2013)
47. McCann, J.: Lessons learned from mondrians applied to real images and color gamuts. In: *Proceedings of Color and Imaging Conference*, vol. 1999, pp. 1–8. Society for Imaging Science and Technology (1999)
48. Mittal, A., Soundararajan, R., Bovik, A.C.: Making a completely blind image quality analyzer. *IEEE Signal Process. Lett.* **20**(3), 209–212 (2012)
49. Morel, J.M., Petro, A.B., Sbert, C.: Fast implementation of color constancy algorithms. In: *Proceedings of Color Imaging XIV: Displaying, Processing, Hardcopy, and Applications*, vol. 7241, pp. 724106–724106. International Society for Optics and Photonics (2009)
50. Morel, J.M., Petro, A.B., Sbert, C.: A pde formalization of retinex theory. *IEEE Trans. Image Process.* **19**(11), 2825–2837 (2010)
51. Ng, M.K., Wang, W.: A total variation model for retinex. *SIAM J. Imaging Sci.* **4**(1), 345–365 (2011)
52. Pallara, L.A.N.F.D., Ambrosio, L., Fusco, N.: Functions of Bounded Variation and Free Discontinuity Problems. Oxford University Press, Oxford (2000)
53. Park, S., Yu, S., Kim, M., Park, K., Paik, J.: Dual autoencoder network for retinex-based low-light image enhancement. *IEEE Access* **6**, 22084–22093 (2018)
54. Parthasarathy, S., Sankaran, P.: An automated multi scale retinex with color restoration for image enhancement. In: *Proceedings of National Conference on Communications (NCC)*, pp. 1–5. IEEE (2012)
55. Provenzi, E., De Carli, L., Alessandro, R., Marini, D.: Mathematical definition and analysis of the retinex algorithm. *JOSA A* **22**(12), 2613–2621 (2005)
56. Provenzi, E., De Carli, L., Rizzi, A., Marini, D.: Mathematical definition and analysis of the retinex algorithm. *JOSA A* **22**(12), 2613–2621 (2005)
57. Provenzi, E., Fierro, M., Rizzi, A., De Carli, L., Gadia, D., Marini, D.: Random spray retinex: a new retinex implementation to investigate the local properties of the model. *IEEE Trans. Image Process.* **16**(1), 162–171 (2006)
58. Rahman, Z.u., Jobson, D.J., Woodell, G.A.: Multi-scale retinex for color image enhancement. In: *Proceedings of 3rd IEEE International Conference on Image Processing*, vol. 3, pp. 1003–1006. IEEE (1996)
59. Rahman, Z., Jobson, D.J., Woodell, G.A.: Retinex processing for automatic image enhancement. *J. Electron. Imaging* **13**(1), 100–111 (2004)
60. Shen, L., Yue, Z., Feng, F., Chen, Q., Liu, S., Ma, J.: Msr-net: Low-light image enhancement using deep convolutional network. arXiv preprint [arXiv:1711.02488](https://arxiv.org/abs/1711.02488) (2017)
61. Shi, W., Loy, C.C., Tang, X.: Deep specialized network for illuminant estimation. In: *European Conference on Computer Vision*, pp. 371–387. Springer (2016)
62. Shi, Y., Wu, X., Zhu, M.: Low-light image enhancement algorithm based on retinex and generative adversarial network. arXiv preprint [arXiv:1906.06027](https://arxiv.org/abs/1906.06027) (2019)
63. Wali, S., Zhang, H., Chang, H., Wu, C.: A new adaptive boosting total generalized variation (tgv) technique for image denoising and inpainting. *J. Vis. Commun. Image Represent.* **59**, 39–51 (2019)
64. Wang, L., Xiao, L., Liu, H., Wei, Z.: Variational bayesian method for retinex. *IEEE Trans. Image Process.* **23**(8), 3381–3396 (2014)
65. Wang, Q., Fu, X., Zhang, X.P., Ding, X.: A fusion-based method for single backlit image enhancement. In: *2016 IEEE International Conference on Image Processing (ICIP)*, pp. 4077–4081. IEEE (2016)
66. Wang, S., Zheng, J., Hu, H.M., Li, B.: Naturalness preserved enhancement algorithm for non-uniform illumination images. *IEEE Trans. Image Process.* **22**(9), 3538–3548 (2013)
67. Wang, W., He, C.: A variational model with barrier functionals for retinex. *SIAM J. Imaging Sci.* **8**(3), 1955–1980 (2015)
68. Wang, W., Li, B., Zheng, J., Xian, S., Wang, J.: A fast multi-scale retinex algorithm for color image enhancement. In: *Proceedings of International Conference on Wavelet Analysis and Pattern Recognition*, vol. 1, pp. 80–85. IEEE (2008)
69. Wang, W., Ng, M.K.: A nonlocal total variation model for image decomposition: illumination and reflectance. *Numer. Math. Theory Methods Appl.* **7**(3), 334–355 (2014)
70. Wei, C., Wang, W., Yang, W., Liu, J.: Deep retinex decomposition for low-light enhancement. arXiv preprint [arXiv:1808.04560](https://arxiv.org/abs/1808.04560) (2018)
71. Zhang, X., Wandell, B.A., et al.: A spatial extension of cielab for digital color image reproduction. In: *Proceedings of SID International Symposium Digest of Technical Papers*, vol. 27, pp. 731–734. Citeseer (1996)

72. Zosso, D., Tran, G., Osher, S.: A unifying retinex model based on non-local differential operators. In: Proceedings of Computational Imaging XI, vol. 8657, pp. 865702–865702. International Society for Optics and Photonics (2013)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Chunxue Wang is an associate researcher at the Cultural Heritage Digitization Institute, Dunhuang Academy, Dunhuang, China. She received her Ph.D. degree (2016) in computational mathematics from the University of Science and Technology of China, China. Her research interests include mathematical image processing, geometry modeling and processing, inverse problems and optimization.



Ligang Liu is a professor at the School of Mathematical Sciences, University of Science and Technology of China. He received his B.Sc. (1996) and his Ph.D. (2001) from Zhejiang University, China. Between 2001 and 2004, he worked at Microsoft Research Asia. Then he worked at Zhejiang University during 2004 and 2012. He paid an academic visit to Harvard University during 2009 and 2011. His research interests include digital geometric processing, computer graphics and image processing. He serves as the associated editor for journals of IEEE TVCG, IEEE CG&A, CGF, CAGD, C&G and The Visual Computer. He served as the conference co-chair of GMP 2017 and the program co-chair of SIAM GD 2019, GMP 2018, CAD/Graphics 2017, CVM 2016, SGP 2015 and SPM 2014. His research works could be found at his research website: <http://sta.ustc.edu.cn/lglui>.



Huayan Zhang is an assistant professor at the School of Computer Science and Technology, Tiangong University, Tianjin, China. She received her Ph.D. (2015) in computational mathematics from the University of Science and Technology of China, China. Her research interests include image processing, geometry modeling and processing.