Manuscript submitted to
AIMS' Journals
Volume **X**, Number **0X**, XX **2019**

doi:10.3934/xx.xx.xx.xx

pp. **1–24**

IMAGE RETINEX BASED ON THE NONCONVEX TV-TYPE REGULARIZATION

Yuan Wang and Zhi-Feng Pang*
School of Mathematics and Statistics
Henan University, Kaifeng, 475004, P.R. China

Yuping Duan
Center for Applied Mathematics
Tianjin University, Tianjin, 300072, P.R. China

Ke Chen
Department of Mathematical Sciences
University of Liverpool, L693BX, United Kingdom

(Communicated by the associate editor name)

Abstract. Retinex theory is introduced to show how the human visual system perceives the color and the illumination effect such as Retinex illusions, medical image intensity inhomogeneity and color shadow effect etc.. Many researchers have studied this ill-posed problem based on the framework of the variation energy functional for decades. However, to the best of our knowledge, the existing models via the sparsity of the image based on the nonconvex ℓ^p-quasinorm were limited. To deal with this problem, this paper considers a TV_p-HOTV_q-based retinex model with $p, q \in (0, 1)$. Specially, the TV_p term based on the total variation(TV) regularization can describe the reflectance efficiently, which has the piecewise constant structure. The HOTV_q term based on the high order total variation(HOTV) regularization can penalize the smooth structure called the illumination. Since the proposed model is non-convex, non-smooth and non-Lipschitz, we employ the iteratively reweighed ℓ_1 (IRL1) algorithm to solve it. We also discuss some properties of our proposed model and algorithm. Experimental experiments on the simulated and real images illustrate the effectiveness and the robustness of our proposed model both visually and quantitatively by compared with some related state-of-the-art variational models.

1. **Introduction.** Our vision tends to see the same color in a given image regardless of the light and the color of objects remains relatively constant under varying illumination. The illusion effect is called the retinex effect. The word Retinex is made up of two words: retina and cortex. The retinex theory was initially proposed by Land and Mccann [27] through modeling the color perception of the human visual system, which can ascertain the reflectance of a object in which both reflectance

2010 *Mathematics Subject Classification.* 80M30, 80M50, 68U10.

Key words and phrases. Image Retinex, TV_p-HOTV_q Regularization, Iteratively Reweighed ℓ_1 Algorithm, Alternating Minimization Method, Bias Field Correction.

Dr. Z.-F. Pang was partially supported by National Basic Research Program of China (973 Program No.2015CB856003), and also gratefully acknowledges financial support from China Scholarship Council(CSC) as a research scholar to visit the University of Liverpool from August 2017 to August 2018.

* Corresponding author: Zhi-Feng Pang. E-mail: wangyuan2028@163.com; zhifengpang@163.com; yuping.duan@tju.edu.cn; K.Chen@liverpool.ac.uk.

and illumination are unknown. In general, we would like to remove the retinex effect to strengthen the image with the nonuniform lighting by separating the illumination from the reflectance in the given image. However, this problem is a classic ill-posed problem due to the lacking of some prior information and the uniqueness.

Many implementations and improvements of the retinex problem have been studied in the literature to overcome the ill-posedness. A path-based algorithm was put forward by Land and Mccann in [27], in which they considered the reflectance at each pixel depending on the multiplication of the ratios along random walks. There were other algorithms based on this approach [36, 46, 8]. The major difference is the path geometry among these path-based algorithms. Piecewise linear paths were used in [28], while double spirals and Brownian paths were used in [8]. But these methods need to tune a mass of parameters and are generally slow in performance. Recursive algorithms were proposed by using a recursive matrix calculation to replace the path computation in [29, 38]. These algorithms are more efficient than a path-based approach. However, it is not clear how many iterations are required in the recursive method and the final results are strongly affected by the number of the iteration setting. Jobson et al. in [22, 23] described the homomorphic filtering scheme for the retinex problem, where they modeled the reflectance as a low-pass filter. Meanwhile, a kernel-based retinex (KBR) [4] was proposed through using the expectation value of a weighted random variable with a suitable kernel function. This scheme can provide an analysis of the action of the KBR on contrast, produce a two-sided contrast modification, and enhance both under- and overexposed picture. Recently, Rizzi [3] and Palma et al. [1] presented other perception based on the color correction in the framework of variational techniques, which allow more flexible local control of the contrast adjustment and attachment to the data.

Recently, the variation-partial differential equation (PDE) based models also got a lot of attention in the image retinex problem, where the models depend on the suitable energy functional space. For the PDE-based models, researchers usually decompose the intensity of the original image into a product of the reflectance and the illumination, where the reflectance and the illumination are usually assumed to be the spatially smoothing and the piecewise constant respectively. With the help of these assumptions, Morel et al. in [40] solved a Poisson equation and Horn in [20] used the Laplacian to the sum $s = l + r$, where s, l, and r are respectively the initial image, illumination, and reflectance in the logarithmic domain, and applied threshold function to clip the peaks of $|\Delta s|$ in order to get a Poisson equation. There are some similar approaches to the work in [5, 39, 40]. However, the numerical methods for these models severely depend on the Courant-Friedrichs-Lewy (CFL) condition when solving the corresponding PDE.

Different to above PDE-based methods, a lot of attention was paid to the variation-based models recently due to the flexible numerical methods and the good mathematical properties based on the choosing functional spaces. Based on the assumption that the illumination is smooth, Kimmel et al. [25] proposed a variational model for explaining how the human vision system perceives colors. However, their model did not consider the reflectance piecewise constant assumption. As one improvement, Ma and Osher [37] coupled the total variation (TV) with the nonlocal TV as the regularization term to the retinex problem and then proposed the Bregman iteration scheme to solve it. But this model did not consider the relationship between the illumination and the reflectance. It was highly hard to present some theoretical analysis such as the existence and the uniqueness of the solution. In

[42], Ng and Wang established a TV model based on the prior of the reflection function, where some constraints are added to make the model more appropriate and reasonable. Liang and Zhang [34] further proposed the higher order TV-based L1 model (called the HOTVL1) for suppressing the staircase artifact generated by the TV regularization [37, 42]. Zosso et al in [53] established a unifying model based on the nonlocal formulation. In addition, Wang and He [50] also proposed a variational model with barrier functionals in order to satisfy some constraints. They solved their model by an alternating minimization and the steepest decent method. By using the L^0 quasi-norm regularization to penalize the sparsity based on some transformations, Duan et al. [12] proposed to combine the Mumford-Shah model with the L^0 quasi-norm for the retinex problem. Gu et al in [18] recently proposed a detail preserving variational model for simultaneously estimating the illumination and the reflectance from an observed image without the log-transform. However, most of the existing variational models perform a log-transform on the original image in their models to turn the product form into the addition form. This operation can reduce the computational complexity and simulate the human eye perception of light intensity. Furthermore, these existing models due to the lacking of the efficient penalty for the noise prior are not robust when the image contains some noises. In order to overcome these drawbacks, Liu et al [32] established a new TV-based model (called the ETV), where the data-fitting term uses the exponent transform to replace the log-transform and one of the regularization terms is used to describe the noise prior and their proposed model can enhance the image details. Aiming to extend this work, this paper employs the non-convex ℓ^p-quasi-norm for improving the ETV and the main contributions are summarized as follows.

- We propose a new image retinex model based on the nonconvex total-variation (TV)-type regularization, where the data fitting term and one of regularization terms based on the exponent-transform are used to describe image details efficiently, one regularization term based on the nonconvex total variation ℓ^p-quasinorm to penalize the sparsity of the piecewise constants and one regularization term based on the nonconvex high-order total variation ℓ^q-quasinorm to penalize the piecewise smoothing regions.
- Since the proposed model is nonconvex, non-smooth and non-Lipschitz, we employ the alternating minimization method, where the iteratively reweighted ℓ^1 algorithm and the alternating direction method of multipliers are used in the sub-minimization problems. We also give some theoretical analysises of the proposed model and the numerical algorithm.
- Numerical comparisons demonstrate that our proposed model achieves competitive results for the image retinex problem with the different types.

The rest of this paper is organized as follows. In section 2, we introduce the iteratively reweighed ℓ_1 algorithm. In section 3, we present the proposed model and some theoretical analysis about the existence of the solution. The efficient numerical methods are proposed to solve the proposed model in this section. Section 4 presents some numerical results to demonstrate the effectiveness of the proposed method compared with some state-of-the-art retinex methods. Finally, we conclude our paper in Section 5.

Notations. For simplification, we set a gray image $f \in \mathbf{X}$. Here $\mathbf{X} =: \mathbb{R}^{M\times N}$ is equipped with the inner product and the norm as

$$\langle \mathbf{x}, \mathbf{y}\rangle = \sum_{i=1}^{M}\sum_{j=1}^{N} x_{i,j}y_{i,j} \text{ and } \|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^{M}\sum_{j=1}^{N} x_{i,j}^2},$$

where $\mathbf{x}, \mathbf{y} \in \mathbf{X}$. Let $\mathbf{Y} := \mathbf{X} \times \mathbf{X}$, the discrete gradient operator $\nabla(\cdot) : \mathbf{X} \to \mathbf{Y}$ is defined by

$$\nabla u_{i,j} = (D_x^+ u_{i,j}, D_y^+ u_{i,j}).$$

Here the forward difference operator D^+ specifically denotes as

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

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

Similarly, we can define the backward difference operator as

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

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

Using above symbols, we get the discrete version of the Hessian operator as

$$\nabla^2 u_{i,j} = (D_x^-(D_x^+ u_{i,j}), D_x^-(D_y^+ u_{i,j}); D_y^-(D_x^+ u_{i,j}), D_y^-(D_y^+ u_{i,j})).$$

In addition, $\forall\ \mathbf{x} := (\mathbf{x}_1, \mathbf{x}_2) \in \mathbf{Y}$ and $\forall\ \mathbf{y} := (\mathbf{y}_1, \mathbf{y}_2) \in \mathbf{Y}$, we define their inner product

$$\langle \mathbf{x}, \mathbf{y}\rangle = \langle \mathbf{x}_1, \mathbf{y}_1\rangle + \langle \mathbf{x}_2, \mathbf{y}_2\rangle.$$

Then the ℓ^1 norm and the ℓ^2-norm are defined in the space $\mathbf{Y}$ as

$$\|\mathbf{x}\|_{2,1} = \sum_{i=1}^{M}\sum_{j=1}^{N} \sqrt{(x_{1,i,j})^2 + (x_{2,i,j})^2} \text{ and } \|\mathbf{x}\|_{2,2} = \sqrt{\langle \mathbf{x}, \mathbf{x}\rangle}.$$

For the simplification, we set $\|\cdot\|_{2,2}$ as $\|\cdot\|_2$ in the following. As an extension, the ℓ^p-quasinorm is defined by

$$\|\mathbf{x}\|_{2,p}^p = \sum_{i=1}^{M}\sum_{j=1}^{N} \left((x_{1,i,j})^2 + (x_{2,i,j})^2\right)^{\frac{p}{2}}$$

for $p \in (0, 1)$. Let $\mathbf{Z} := \mathbf{Y} \times \mathbf{Y}$ and $\forall\ \boldsymbol{\rho} = (\rho_{11}, \rho_{12}; \rho_{21}, \rho_{22}) \in \mathbf{Z}$, the ℓ^q-quasinorm based on the space $\mathbf{Z}$ is defined by

$$\|\boldsymbol{\rho}\|_{\mathrm{F},q}^q = \sum_{i=1}^{M}\sum_{j=1}^{N} \left((\rho_{11,i,j})^2 + (\rho_{12,i,j})^2 + (\rho_{21,i,j})^2 + (\rho_{22,i,j})^2\right)^{\frac{q}{2}}$$

for $q \in (0, 1)$. Here F denotes the Frobenius norm. Let div and div^2 to denote the adjoint operators of ∇ and ∇^2 respectively, based on the divergence theorem

$$\langle -\mathrm{div}\mathbf{x}, u\rangle_{\mathbf{X}} = \langle \mathbf{x}, \nabla u\rangle_{\mathbf{Y}} \text{ and } \langle \mathrm{div}^2\boldsymbol{\rho}, u\rangle_{\mathbf{X}} = \langle \boldsymbol{\rho}, \nabla^2 u\rangle_{\mathbf{Z}},$$

we have

$$\begin{aligned}
&\mathrm{div}\mathbf{x}_{i,j} = D_x^- x_{1,i,j} + D_y^- x_{2,i,j},\\
&\mathrm{div}^2\boldsymbol{\rho}_{i,j} = D_x^-\left(D_x^+\boldsymbol{\rho}_{11,i,j}\right) + D_x^-\left(D_y^+\boldsymbol{\rho}_{12,i,j}\right) + D_y^-\left(D_x^+\boldsymbol{\rho}_{21,i,j}\right) + D_y^-\left(D_y^+\boldsymbol{\rho}_{22,i,j}\right).
\end{aligned}$$

2. **Iteratively reweighted l_1 algorithm.** This section recalls some numerical methods related to our proposed model and the numerical algorithm.

2.1. **The $\ell^2-\ell^p$ proximal problem.** During last decades, how to solve the $\ell^2-\ell^p$ proximal problem, as one subproblem used in the field of compressed sensing, pattern recognition and image restoration, has attracted great attention [7, 9, 54, 33, 31]. Formally, this problem can be written as

$$\min_{\mathbf{x}} \frac{\tau}{2}\|\mathbf{x}-\mathbf{y}\|_{\mathfrak{X}}^2 + \|\mathbf{x}\|_{\mathfrak{X},p}^p, \tag{1}$$

where $p\in(0,1)$ and $\mathfrak{X}$ depends on some spaces such as $\mathbf{Y}$ and $\mathbf{Z}$. That is to say, we set $\mathfrak{X}:=2$ if $\mathbf{x}\in\mathbf{X}$ and set $\mathfrak{X}:=\mathrm{F}$ if $\mathbf{y}\in\mathbf{Y}$. Since the numerical method based on a different space in $\mathfrak{X}$ is similar, we in the following consider the case of $\mathbf{x}\in\mathbf{X}$ as

$$\min_{\mathbf{x}} \frac{\tau}{2}\|\mathbf{x}-\mathbf{y}\|_2^2 + \|\mathbf{x}\|_{2,p}^p. \tag{2}$$

To solve the problem (2), there are some great challenges for the theoretical analysis and the numerical implementation due to its non-convexity, non-smoothing and non-Lipschitz. Many schemes have been proposed to overcome these drawbacks such as the smoothing approximation method [9], the general iterative shrinkage-thresholding algorithm [54, 31], the active set method [15] and the iterative reweighed ℓ^1 minimization algorithm (IRLA) [7, 26, 35]. Here we mainly use the latter to solve the problem (2).

The IRLA was first proposed by Candes et al. [7] to slove the problem (2). This algorithm consists of solving a sequence of the weighted ℓ^1-minimization problems, where the weighted parameter used for the next iteration are computed from the value of the current solution. Specifically, solving the original problem (2) is approximately replaced by

$$\mathbf{x}_{i,j}^{k+1} = \underset{\mathbf{x}_{i,j}}{\mathrm{argmin}} \underbrace{\frac{\tau}{2}\|\mathbf{x}_{i,j}-\mathbf{y}_{i,j}\|_2^2 + T_p(\mathbf{x}_{i,j}^k,\eta^k)\|\mathbf{x}_{i,j}\|_{2,1}}_{\mathcal{G}\left(\mathbf{x}_{i,j},\mathbf{x}_{i,j}^k,\eta^k\right)} \tag{3}$$

for $i=1\cdots M$ and $j=1\cdots N$, where $\mathbf{x}_{i,j}:=(x_{1,i,j},x_{2,i,j})$ and $T_p\left(\mathbf{x}_{i,j}^k,\eta^k\right):=\left(\left\|\mathbf{x}_{i,j}^k\right\|_{2,1}+\eta^k\right)^{p-1}$. It is obvious that this problem is the classical $\ell^2-\ell^1$ problem and the closed-form solution can be obtained by using the soft shrinkage scheme

$$\mathbf{x}_{i,j}^{k+1} = \mathrm{shrink}\left(\mathbf{y}_{i,j},\frac{T_p\left(\mathbf{x}_{i,j}^k,\eta^k\right)}{\tau}\right) := \max\left\{\|\mathbf{y}_{i,j}\|_{2,1} - \frac{T_p\left(\mathbf{x}_{i,j}^k,\eta^k\right)}{\tau},0\right\}\frac{\mathbf{y}_{i,j}}{\|\mathbf{y}_{i,j}\|_{2,1}}.$$

In the following, we summarize the IRLA to solve the problem (2) as Algorithm 1.

Algorithm 1: Iteratively Reweighed l_1 Algorithm for the problem (2)

1. **Parameter**: $\varepsilon > 0$ and $\tau > 0$;
2. **for** $k = 1, 2, \cdots\cdots$, **do**

$$\begin{cases} \mathbf{x}_{i,j}^{k+1} = \text{shrink}\left(\mathbf{y}_{i,j}, \frac{T\left(\mathbf{x}_{i,j}^{k}, \eta^k, p\right)}{\tau}\right); \\ \eta^{k+1} = \rho\eta^k \end{cases}$$

   for $i = 1 \cdots M$, $i = 1 \cdots N$ and $\rho \in (0, 1)$.
3. **End while** the stopping conditions are satisfied.

If setting $\rho = 1$ in Algorithm 1, that is to say, we use the fixed smooth parameter η in the approximated problem (3), the convergence analysis has been discussed in [10]. However, the generated sequence $\{\mathbf{x}^k\}$ only converges to the solution of the smooth problem (3) rather than the solution of the original problem (2). In order to give the convergence analysis of Algorithm 1, we need to show the existence of the solution for the problem (2).

Theorem 2.1. *$\forall\ \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}$ and set $\tau > 0$, the problem (2) exists one solution.*

In fact, the objective function in the problem (2) is bounded from below and coercive, it is easy to deduce the existence of the solution in the problem (2) as done in [15]. Now we consider the convergence of Algorithm 1 related to the framework in [24].

Theorem 2.2. *The sequence $\{\boldsymbol{x}^n\}$ generated by Algorithm 1 converges to one solution of the problem (2).*

Proof. Since the problem (2) is separable, we only consider the case at the pixel point (i, j). Based on the problem (3), we can deduce that

$$\mathcal{G}\left(\mathbf{x}_{i,j}^{k+1}, \mathbf{x}_{i,j}^{k}, \eta^k\right) \leq \mathcal{G}\left(\mathbf{x}_{i,j}^{k}, \mathbf{x}_{i,j}^{k}, \eta^k\right). \tag{4}$$

Omitting the iteration superscript in the problem (3), we can rewrite $\mathcal{G}\left(\mathbf{x}_{i,j}, \mathbf{x}_{i,j}, \eta\right)$ as

$$\begin{aligned} \mathcal{G}\left(\mathbf{x}_{i,j}, \mathbf{x}_{i,j}, \eta\right) &= \frac{\tau}{2}\|\mathbf{x}_{i,j} - \mathbf{y}_{i,j}\|_2^2 + T_p(\mathbf{x}_{i,j}, \eta)\|\mathbf{x}_{i,j}\|_{2,1} \\ &= \frac{\tau}{2}\|\mathbf{x}_{i,j} - \mathbf{y}_{i,j}\|_2^2 + p(\|\mathbf{x}_{i,j}\|_{2,1} + \eta)T_p(\mathbf{x}_{i,j}, \eta) + (1-p)T_p(\mathbf{x}_{i,j}, \eta)^{\frac{p}{p-1}}. \end{aligned}$$

Set $\mathbf{x}_{i,j} := \mathbf{x}_{i,j}^{k+1}$ in the norm $\|\cdot\|_2$ and $\eta := \eta^k$ in above equation, it is easy to get

$$\mathbf{x}_{i,j}^{k+1} = \underset{\mathbf{x}_{i,j}}{\text{argmin}}\, \mathcal{G}\left(\mathbf{x}_{i,j}^{k+1}, \mathbf{x}_{i,j}, \eta^k\right).$$

Combining it with the inequation (4), we have

$$\mathcal{G}\left(\mathbf{x}_{i,j}^{k+1}, \mathbf{x}_{i,j}^{k+1}, \eta^k\right) \leq \mathcal{G}\left(\mathbf{x}_{i,j}^{k+1}, \mathbf{x}_{i,j}^{k}, \eta^k\right) \leq \mathcal{G}\left(\mathbf{x}_{i,j}^{k}, \mathbf{x}_{i,j}^{k}, \eta^k\right). \tag{5}$$

In addition, since $\eta^{k+1} \leq \eta$, we deduce

$$T_p(\mathbf{x}_{i,j}^{k+1}, \eta^{k+1}) \leq T_p(\mathbf{x}_{i,j}^{k+1}, \eta^k).$$

and then obtain that

$$\mathcal{G}\left(\mathbf{x}_{i,j}^{k+1}, \mathbf{x}_{i,j}^{k+1}, \eta^{k+1}\right) \leq \mathcal{G}\left(\mathbf{x}_{i,j}^{k+1}, \mathbf{x}_{i,j}^{k+1}, \eta^k\right) \leq \mathcal{G}\left(\mathbf{x}_{i,j}^{k+1}, \mathbf{x}_{i,j}^{k}, \eta^k\right) \leq \mathcal{G}\left(\mathbf{x}_{i,j}^{k}, \mathbf{x}_{i,j}^{k}, \eta^k\right).$$

This result illustrates that the sequence $\left\{\mathcal{G}\left(\mathbf{x}_{i,j}^k, \mathbf{x}_{i,j}^k, \eta^k\right)\right\}$ is non-increasing. Based on the bounded blow of the problem (3) and $\eta^k \to 0$ as $k \to \infty$, we can conclude that the sequence $\left\{\mathcal{G}\left(\mathbf{x}_{i,j}^k, \mathbf{x}_{i,j}^k, \eta^k\right)\right\}$ has a accumulation as

$$\lim_{k\to+\infty} \mathcal{G}\left(\mathbf{x}_{i,j}^k, \mathbf{x}_{i,j}^k, \eta^k\right) = \mathcal{G}(\bar{\mathbf{x}}_{i,j}, \bar{\mathbf{x}}_{i,j}, 0).$$

for $\bar{\mathbf{x}}_{i,j}^k \to \bar{\mathbf{x}}_{i,j}$ as $k \to \infty$ [1] Now we only want to show that $\bar{\mathbf{x}}_{i,j}$ is one solution of the problem (3). In fact, we plug the point $(\bar{\mathbf{x}}_{i,j}, \bar{\mathbf{x}}_{i,j}, 0)$ in the problem (4), it is easy to find that the problem (4) is equivalent to the problem (3) for $i = 1, \cdots\cdots M$ and $j = 1, \cdots\cdots N$. □

Recently, Ochs et al. [43] extended it to solve various kinds of iteratively reweighed algorithms for the computer vision problems.

3. Total variation model.

In this section, we first introduce our proposed model to the retinex problem and then give an efficient numerical method to solve it.

3.1. The proposed model.

The intensity in the observed image is determined by two parts: the intensity of the illumination received at this point and the ratio of the illumination reflected from this point [21]. Assuming that the illumination of one object is inhomogeneous, it is natural to model the observation through inducing the varying illumination [30]. The varying illumination may have many different sources such as the varying distance based on a point lighting source, the uneven thickness of clouds via filtering sun lights, or the biased magnetic field generated in an MRI machine. Regardless of the sources, we furthermore assume that the illumination should vary smoothly. For simplification, this paper studies and discusses our proposed model based on a single channel-like gray-level image. Specially, the observation image f can be modeled by:

$$f = R \odot L + \eta \tag{6}$$

where $\odot$ denotes the entrywise multiplication and η denotes the additive noise. Here we need two assumptions, which include that the source reflectance $R \in (0, 1)$ is the piecewise constant and the illumination $L \in (0, \infty)$ is the spatially smoothing.

In the model (6), the multiplication operation leads to many numerical difficulties to obtain R and L. One efficient method is of decoupling with them by using the logarithm transformation as

$$v = r + l,$$

where $r = \log(R)$ and $l = \log(L)$. With this manipulation, the model (6) can be rewritten as

$$f = e^{r+l} + \eta. \tag{7}$$

The problem (7) is the classic additive model and many schemes were proposed to deal with it based on some prior assumptions. Here we assume that η is the white Gaussian noise and then consider the following retinex model

$$\begin{aligned}\min_{v,r,l} \mathbb{F}(v, r, l) \quad = \quad & \frac{1}{2}\left\|f - e^v\right\|_2^2 + \alpha\left\|\nabla e^v\right\|_{2,1} + \beta\|\nabla r\|_{2,p}^p + \gamma\left\|\nabla^2 l\right\|_{\mathrm{F},q}^q \\ & + \frac{\mu}{2}\|v - r - l\|_2^2 + \frac{\theta}{2}\|l\|_2^2,\end{aligned} \tag{8}$$

[1] Since the sequence $\left\{\mathcal{G}\left(\mathbf{x}_{i,j}^k, \mathbf{x}_{i,j}^k, \eta^k\right)\right\}$ has a accumulation, we only set $\mathbf{x}_{i,j} = \bar{\mathbf{x}}_{i,j}$ to get the value of this accumulation for the functional $G(\cdot)$.

where α, β and γ are the positive regularization parameter, μ is the penalty parameter. The term $\|f - e^v\|_2^2$ is used for the fidelity, $\|\nabla e^v\|_1$, $\|\nabla r\|_{2,p}^p$ and $\|\nabla^2 l\|_{\mathrm{F}\mathbf{F},q}^q$ are the regularization term, and $\|v - r - l\|_2^2$ is the penalty term. The last term here is to ensure the well-posedness of the model, where θ can be taken to be very small.

We remark the difference between the ETV used in [32] and the proposed model (8). It is obvious that the model (8) is the ETV if setting $p = q = 1$. However, as is well known that the image structure will be sparse based on some transformations such as the gradient-based operators and the wavelet-based operators. So it is more reasonable for restricting parameters $p, q \in (0, 1)$ in order to preserve the sparsity when proposing the reconstruction model as done in (8). This fact can be observed that our proposed model (8) can give better results in the quality of the solution compared with the ETV from the numerical comparisons.

Theorem 3.1. *Let the parameters α, β, γ, μ and θ to be positive, then the problem (8) exists one solution in the space $\mathbb{BV}(\boldsymbol{X}) \times \mathbb{BV}_{\mathrm{p}}(\boldsymbol{X}) \times \mathbb{HBV}_{\mathrm{q}}(\boldsymbol{X})$[2].*

Proof. Since $\mathbb{F}(v, r, l)$ is bounded from below, the existence can be followed from its continuity and coercivity. Let $\mathbf{w} := (v, r, l) = \left(s_1\bar{v}, s_2\bar{r}, s_3\bar{l}\right) \in \mathbb{BV}(\mathbf{X}) \times \mathbb{BV}_{\mathrm{p}}(\mathbf{X}) \times \mathbb{HBV}_{\mathrm{q}}(\mathbf{X})$ such that $s_i > 0$ and $\|\bar{v}\|_2 = 1$, $\|\bar{r}\|_2 = 1$, $\left\|\bar{l}\right\|_2 = 1$. We first notice that

$$\mathbb{F}(\mathbf{w}) \geq \beta s_1 \|\nabla \bar{r}\|_{2,p}^p + \gamma s_2 \left\|\nabla^2 \bar{l}\right\|_{\mathrm{F},q}^q + s_3 \frac{\theta}{2} \|\bar{l}\|_2^2 \to \infty$$

when $\mathbf{w} \to \infty$, that is to say, $s_i \to +\infty$ for $i = 1, 2, 3$. So $\mathbb{F}(\mathbf{w})$ is coercive and then the minimizing sequence $\left\{\mathbf{w}^k\right\}$ for the functional $\mathbb{F}\left(\mathbf{w}^k\right)$ must be bounded. Boundary of the sequence in the space $\mathbf{X}$ implies the existence of a convergence subsequence of $\left\{\mathbf{w}^{k_t}\right\}$. Using the continuity of the function $\mathbb{F}(\mathbf{w})$ in the closed image domain, above assertion is held. □

3.2. **The alternating minimization algorithm.** The problem (8) is a multivariable optimization problem, where the variables v, r, l are coupled together. In order to decouple with them, one efficient method is of using the alternating minimization method to transforming it into some easy and solvable subproblems. In general, the alternating minimization (AM) method, which is a rather old and fundamental algorithm, is attractive due to its simplicity and efficiency. This method is widely used in the linearly constrained optimization problems such as the machine learning, the image and signal processing, as well as other fields [17]. Specially, we can summarize this method as Algorithm 2 in the following.

Algorithm 2: Alternating minimization algorithm for the problem (8)

1. Initialization: Set $v^0 = r^0 = \log(f)$ and parameters α, β, γ, μ, θ;
2. Iterations:

$$\begin{cases} v^{n+1} = \underset{v}{\operatorname{argmin}} \ \dfrac{1}{2}\|f - e^v\|_2^2 + \alpha\|\nabla e^v\|_{2,1} + \dfrac{\mu}{2}\|v - r^n - l^n\|_2^2; & (9) \\ r^{n+1} = \underset{r}{\operatorname{argmin}} \ \beta\|\nabla r\|_{2,p}^p + \dfrac{\mu}{2}\|v^{n+1} - r - l^n\|_2^2; & (10) \\ l^{n+1} = \underset{l}{\operatorname{argmin}} \ \gamma\|\nabla^2 l\|_{\mathrm{F},q}^q + \dfrac{\mu}{2}\|v^{n+1} - r^{n+1} - l\|_2^2 + \dfrac{\theta}{2}\|l\|_2^2; & (11) \end{cases}$$

[2] $\forall \mathbf{w} \in \mathbf{X}$, the bounded TV-type spaces are defined by $\mathbb{BV}(\mathbf{X}) := \{w | \|\nabla w\|_{2,1}\} \leq \infty$, $\mathbb{BV}_{\mathrm{p}}(\mathbf{X}) := \left\{w | \|\nabla w\|_{2,p}^p\right\} \leq \infty$, $\mathbb{HBV}_{\mathrm{q}}(\mathbf{X}) := \left\{w | \left\|\nabla^2 w\right\|_{\mathrm{F},q}^q\right\} \leq \infty$

3. End till the stopping criterion meets:

$$\max(R(r^n, l^n)) \leq 10^{-5} \text{ or } n \geq 5 \times 10^3,$$

where the $R(r^n, l^n)$ is defined in Section 4.

Based on Algorithm 2, we can find that the minimization sequence $\{\mathbb{F}(v^n, r^n, l^n)\}$ is monotonous and nonincreasing. Then we can deduce that there exists a subsequence of $\{\mathbb{F}(v^n, r^n, l^n)\}$ to converge to an accumulation point of the problem (8). In the following, we consider how to solve every subproblem in Algorithm 2.

3.2.1. ***Subproblem*** *(9).* This subproblem is nonconvex and nonsmooth due to the existence of the nonlinear function e^v and the ℓ^1-norm. One efficient method is of using the alternating direction method of multipliers (ADMM) through introducing some auxiliary variables to break the original problem into smaller pieces, each of which are then easier to be handled [17, 2, 13, 11, 6, 16, 47, 51]. If the original problem is convex, the ADMM is closely related to many other methods such as the dual decomposition, the method of multipliers, Douglas-Rachford splitting, Spingarns method of partial inverses, Dykstras alternating projections; see for the work in [17]. To this end, we need to first introduce two auxiliary variables u, $\mathbf{b}$ and then transform the problem (9) into an equivalent form as

$$\begin{cases} \min\limits_{v,u,\mathbf{b}} \dfrac{1}{2}\|f-u\|_2^2 + \alpha\|\mathbf{b}\|_{2,1} + \dfrac{\mu}{2}\|v - r^n - l^n\|_2^2 \\ \text{s.t. } u = e^v, \ \mathbf{b} = \nabla u. \end{cases} \tag{12}$$

Based on the framework of the augmented Lagrangian method, we need to rewrite (12) as the equivalent unconstrained optimization problem

$$\begin{aligned} \max_{\Lambda_1, \mathbf{\Lambda_2}} \min_{v,u,\mathbf{b}} \frac{1}{2}\|f-u\|_2^2 + \alpha\|\mathbf{b}\|_{2,1} + \frac{\mu}{2}\|v - r^n - l^n\|_2^2 + \langle \Lambda_1, u - e^v\rangle \\ + \frac{r_1}{2}\|u - e^v\|_2^2 + \langle \mathbf{\Lambda_2}, \mathbf{b} - \nabla u\rangle + \frac{r_2}{2}\|\mathbf{b} - \nabla u\|_2^2, \end{aligned} \tag{13}$$

where Λ_1 and $\mathbf{\Lambda}_2$ are the Lagrange multiplier, r_1 and r_2 are the positive penalty parameter. Using the ADMM, we can solve the saddle point problem (13) by the following Gaussian-Seidel iteration scheme

$$\begin{cases} v^{k+1} = \operatorname*{argmin}\limits_{v} \dfrac{\mu}{2}\|v - r^n - l^n\|_2^2 - \langle \Lambda_1^k, e^v\rangle + \dfrac{r_1}{2}\|u^k - e^v\|_2^2, & (14) \\ u^{k+1} = \operatorname*{argmin}\limits_{u} \dfrac{1}{2}\|f-u\|_2^2 + \langle \Lambda_1^k, u\rangle + \dfrac{r_1}{2}\|u - e^{v^{k+1}}\|_2^2 \\ \qquad\qquad -\langle \mathbf{\Lambda_2}^k, \nabla u\rangle + \dfrac{r_2}{2}\|\mathbf{b}^k - \nabla u\|_2^2, & (15) \\ \mathbf{b}^{k+1} = \operatorname*{argmin}\limits_{\mathbf{b}} \alpha\|\mathbf{b}\|_{2,1} + \langle \mathbf{\Lambda_2}^k, \mathbf{b}\rangle + \dfrac{r_2}{2}\|\mathbf{b} - \nabla u^{k+1}\|_2^2, & (16) \\ \Lambda_1^{k+1} = \Lambda_1^k + r_1\left(u^{k+1} - e^{v^{k+1}}\right), & (17) \\ \mathbf{\Lambda}_2^{k+1} = \mathbf{\Lambda}_2^k + r_2\left(\mathbf{b}^{k+1} - \nabla u^{k+1}\right). & (18) \end{cases}$$

In the iteration scheme (14)-(18), it is not easy to solve the subproblem (14) since the nonlinear function e^v makes the objective function be nonconvex. So we either use an approximate solution as a substitute in the update which might cause divergence, or solve the subproblems by numerical algorithms which can bring

computational burden. One efficient method is of using the Taylor linearizations around and then adds a proximal term as

$$v^{k+1} = \underset{v}{\operatorname{argmin}} \frac{\mu}{2}\|v - r^n - l^n\|_2^2 + \left(r_1 e^{v^k}\left(e^{v^k} - u^k\right) - e^{v^k}\Lambda_1^k\right) v + \frac{\tau}{2}\left\|v - v^k\right\|_2^2.$$

This is a smooth and convex optimization problem, the optimal condition of which can be obtained by

$$(\tau + \mu)v^{k+1} = \tau v^k + e^{v^k}\Lambda_1^k - r_1 e^{v^k}\left(e^{v^k} - u^k\right) + \mu(r^n + l^n).$$

Furthermore, we have

$$v^{k+1} = \frac{\tau v^k + e^{v^k}\Lambda_1^k - r_1 e^{v^k}\left(e^{v^k} - u^k\right) + \mu(r^n + l^n)}{\tau + \mu} \tag{19}$$

To the subproblem (15), it is a convex and smooth. With the help of the variational method and simultaneously choosing the Neumann boundary condition, we can get the corresponding Euler-Lagrangian equation as follows

$$((1 + r_1)\mathcal{I} - r_2\Delta)u = f - \Lambda_1^k + r_1 e^{v^{k+1}} - \operatorname{div}\mathbf{\Lambda_2^k} - r_2\operatorname{div}\mathbf{b}^k, \tag{20}$$

where $\mathcal{I}$ is the identity operator and Δ is the Laplace operator. This problem is of solving the linear equation, and the numerical method depends on the boundary condition [52]. This paper mainly uses the periodic boundary condition. So the matrix operator $(1 + r_1)\mathcal{I} - r_2\Delta$ is a circulant matrix and the solution can be obtained by using the fast Fourier transform (FFT):

$$u^{k+1} = \mathcal{F}^{-1}\left(\frac{\mathcal{F}\left(f - \Lambda_1^k + r_1 e^{v^{k+1}}\right) - \mathcal{F}(\operatorname{div}\mathbf{\Lambda_2^k}) - r_2\mathcal{F}(\operatorname{div}\mathbf{b}^k)}{(1 + r_1)\mathcal{F}(\mathcal{I}) - r_2\mathcal{F}(\Delta)}\right), \tag{21}$$

where $\mathcal{F}$ denotes the Fourier transform and $\mathcal{F}^{-1}$ denotes its inverse transform.

To solve the subproblem (16), it corresponds to the classical $\ell^2 - \ell^1$ problem and the closed-form solution can be obtained by using the soft-shrinkage operator as

$$\begin{aligned}\mathbf{b}^{k+1} &= \underset{\mathbf{b}}{\operatorname{argmin}}\, \alpha\|\mathbf{b}\|_{2,1} + \frac{r_2}{2}\left\|\mathbf{b} - \left(\nabla u^{k+1} - \frac{\mathbf{\Lambda_2^k}}{r_2}\right)\right\|_2^2 \\ &= \operatorname{shrink}\left(\nabla u^{k+1} - \frac{\mathbf{\Lambda_2^k}}{r_2}, \frac{\alpha}{r_2}\right).\end{aligned}$$

3.2.2. ***Subproblem*** *(10)*. It is not easy to solve this problem due to its nonconvexity, nonsmoothing and non-Lipschitz. In order to overcome these numerical difficulties, we need to use the operator splitting scheme by introducing an auxiliary variable $\mathbf{w} := \nabla r$ and then obtain a constrained optimization problem

$$\begin{cases}\min\limits_{r,\mathbf{w}} \frac{\mu}{2}\left\|v^{n+1} - r - l^n\right\|_2^2 + \beta\|\mathbf{w}\|_{2,p}^p, \\ \text{s.t. } \mathbf{w} = \nabla r,\end{cases}$$

where $v^{n+1} := v^{k+1}$. Here v^{k+1} denotes the outputting of the inner iteration related to the subproblem (9). Using the augmented Lagrangian method, we can obtain a saddle point problem as

$$\max_{\mathbf{\Lambda_3}} \min_{r,\mathbf{w}} \frac{\mu}{2}\left\|v^{n+1} - r - l^n\right\|_2^2 + \beta\|\mathbf{w}\|_{2,p}^p + \langle\mathbf{\Lambda_3}, \mathbf{w} - \nabla r\rangle + \frac{r_3}{2}\|\mathbf{w} - \nabla r\|_2^2, \tag{22}$$

where $\mathbf{\Lambda_3}$ is a Lagrange multiplier, and r_3 is a positive penalty parameter. Based on the ADMM, we can solve the problem (22) through the following iteration scheme

$$\begin{cases} r^{\kappa+1} = \underset{r}{\operatorname{argmin}} \frac{\mu}{2}\|v^{n+1} - r - l^n\|_2^2 - \langle \mathbf{\Lambda_3^\kappa}, \nabla r\rangle + \frac{r_3}{2}\left\|\mathbf{w}^\kappa - \nabla r\right\|_2^2, & (23) \\ \mathbf{w}^{\kappa+1} = \underset{\mathbf{w}}{\operatorname{argmin}}\, \beta\|\mathbf{w}\|_{2,p}^p + \langle \mathbf{\Lambda_3^\kappa}, \mathbf{w}\rangle + \frac{r_3}{2}\left\|\mathbf{w} - \nabla r^{\kappa+1}\right\|_2^2, & (24) \\ \mathbf{\Lambda_3^{\kappa+1}} = \mathbf{\Lambda_3^\kappa} + r_3\left(\mathbf{w}^{\kappa+1} - \nabla r^{\kappa+1}\right). & (25) \end{cases}$$

Now we consider to solve the smooth subproblem (23), the optimal condition of which can be written as

$$(\mu\mathcal{I} - r_3\Delta)r = \mu\left(v^{n+1} - l^n\right) - r_3\mathrm{div}\mathbf{w}^\kappa - \mathrm{div}\mathbf{\Lambda_3^\kappa}. \tag{26}$$

Based on the assumption of the periodic boundary condition again, the solution

$$r^{\kappa+1} = \mathcal{F}^{-1}\left(\frac{\mu\mathcal{F}\left(v^{n+1} - l^n\right) - \mathcal{F}(\mathrm{div}\mathbf{\Lambda_3^\kappa}) - r_3\mathcal{F}\left(\mathrm{div}\mathbf{w}^\kappa\right)}{\mu\mathcal{F}(\mathcal{I}) - r_3\mathcal{F}(\Delta)}\right) \tag{27}$$

of this problem can be obtained by using the FFT.

To the problem (24), we rewrite it as

$$\min_{\mathbf{w}} \beta\|\mathbf{w}\|_{2,p}^p + \frac{r_3}{2}\left\|\mathbf{w} - \left(\nabla r^{\kappa+1} - \frac{\mathbf{\Lambda_3^\kappa}}{r_3}\right)\right\|_2^2 \tag{28}$$

and then employ Algorithm 1 to solve it by setting

$$\mathbf{y} := \nabla r^{\kappa+1} - \frac{\mathbf{\Lambda_3^\kappa}}{r_3} \text{ and } \tau = \frac{r_3}{\beta}$$

in the problem (1).

3.2.3. ***Subproblem*** *(11).* The method for solving this problem is similar to solving the problem (10). So we similarly introduce an auxiliary variable $\mathbf{z} := \nabla^2 l$ and use the ADMM to solve it as follows.

$$\begin{cases} l^{\nu+1} = \underset{l}{\operatorname{argmin}} \frac{\mu}{2}\|v^{n+1} - r^{n+1} - l\|_2^2 + \frac{\theta}{2}\|l\|_2^2 - \langle \mathbf{\Lambda_4^\nu}, \nabla^2 l\rangle + \frac{r_4}{2}\|\mathbf{z}^\nu - \nabla^2 l\|_2^2, & (29) \\ \mathbf{z}^{\nu+1} = \underset{\mathbf{z}}{\operatorname{argmin}}\, \gamma\|\mathbf{z}\|_{\mathrm{F},q}^q + \langle \mathbf{\Lambda_4^\nu}, \mathbf{z}\rangle + \frac{r_4}{2}\|\mathbf{z} - \nabla^2 l^{\nu+1}\|_2^2, & (30) \\ \mathbf{\Lambda_4^{\nu+1}} = \mathbf{\Lambda_4^\nu} + r_4\left(\mathbf{z}^{\nu+1} - \nabla^2 l^{\nu+1}\right). & (31) \end{cases}$$

where $v^{n+1} := v^{k+1}$ and $r^{n+1} := r^{\kappa+1}$. Here v^{k+1} and $r^{\kappa+1}$ denote the outputting of the inner iteration related to the subproblem (9) and (10).

In the iteration scheme (29)-(31), the subproblem (29) is smooth and convex. Then the optimization condition can be written as

$$\left((\mu+\theta)\mathcal{I} + r_4\mathrm{div}^2\nabla^2\right) l = \mu\left(v^{n+1} - r^{n+1}\right) + r_4\mathrm{div}^2\mathbf{z}^\omega + \mathrm{div}^2\mathbf{\Lambda_4^\omega}.$$

Using the assumption of the periodic boundary condition, the solution can be obtained by using the FFT through

$$l^{\omega+1} = \mathcal{F}^{-1}\left(\frac{\omega\mathcal{F}\left(v^{n+1} - r^{n+1}\right) + \mathcal{F}\left(\mathrm{div}^2\mathbf{\Lambda_4^\omega}\right) + r_4\mathcal{F}\left(\mathrm{div}^2\mathbf{z}^\omega\right)}{(\mu+\theta)\mathcal{F}(\mathcal{I}) + r_4\mathcal{F}\left(\mathrm{div}^2\nabla^2\right)}\right). \tag{32}$$

Similarly, the subproblem (30) can be written as

$$\min_{\mathbf{z}} \gamma\|\mathbf{z}\|_{\mathrm{F},q}^q + \frac{r_4}{2}\left\|\mathbf{z} - \left(\nabla^2 l^{\omega+1} - \frac{\mathbf{\Lambda_4^\omega}}{r_4}\right)\right\|_2^2, \tag{33}$$

which is the $\ell^2 - \ell^q$ problem. Then we still use Algorithm 1 to solve it by setting

$$\mathbf{y} := \nabla^2 l^{\omega+1} - \frac{\mathbf{\Lambda_4^\omega}}{r_4} \text{ and } \tau = \frac{r_4}{\gamma}$$

in the problem (1).

Remark 1. To Algorithm 2, there include some inner iterations to solve the subproblem (9)-(11). Specifically, we use the ADMM to solve them respectively. In general, more exact solution can be obtained when increasing the number of iteration. However, we notice that a better decomposition result can be still obtained by setting the inner iteration to be one. So we choose it as the stopping condition for all of the inner iterations in the numerical implementation when solving our proposed model.

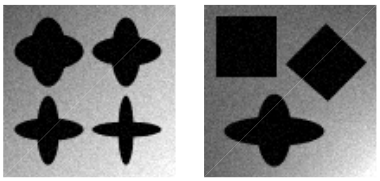
4. **Numerical experiments.** To demonstrate the performance of our proposed model (8), this section presents a series of numerical implementations to remove the different illusions and noisy levels of the different kind of the images. To this end, we choose some state-of-the art models such as the TVH1 [42], the HOTVL1 [34], the L0MS [12] and the ETV [32] compared with our proposed model (8).

As is well known, choosing suitable parameter in the proposed model is essential for obtaining a good correction and the reconstruction result. For the numerical algorithm, the suitable parameters can improve the convergence rate and the numerical stability. However, there involve more than two parameters such as the regularization parameters α, β and γ, the penalty parameters μ, p and q. In general, three ways can be proposed for how to select them. One is of trying on some values for providing satisfactory results and then keeping it to be the fixed outputting. The second way is by using Game theory to reformulate the optimization problem so that the dimension of the parameters' space is much reduced, as is done in [48]. The third approach is of choosing them by using some rules based on the prior statistical characteristic such as the L-curve method [19], the generalized cross validation (GCV) [14, 49], the discrepancy principle [41, 45], or the variational Bayes approach [44]. In our experiments, we choose the former approach. Specifically, we set $\alpha \in [0.01, 0.04]$ according to the severity of noises, $\beta \in [0.001, 0.02]$ depending on the illusion, $\mu \in [0.7, 1.2]$ and $p \in [0.9, 0.97]$ based on the image structure. In addition, we fix $\gamma = 0.08, \tau = 1e-5, q = 0.85$ in the numerical implementation. In addition, in order to quantitate the performance, we adopt the Peak Signal to Noise Ratio (PSNR) and the Mean Structural Similarity (MSSIM) to evaluate the effectiveness of the used models. All comparison algorithms are implemented and tested in Matalb R2018a under Windows OS on a laptop computer equipped with Intel Core i7 2.7GHz.

4.1. **Synthetic Images.** We first start with reconstructing two synthetic images degraded by the white Gaussian noise with $\sigma = 0.001$ and a suitable bias field as shown in Figure 1. From the values of the PSNR and the MSSIM in Table 1, we can conclude that our method demonstrates competitive score. Since humans are involved in evaluating the quality of the reconstructed result, it absolutely becomes essential to provide visual quality of the reconstructed images. As shown in Figure 2, we can observe that the recovered reflectance and illumination images generated by our proposed model are less noise effects and artifacts than other models. Especially, some noises can be found in Figure 2 (a), (b) and (c) and some piecewise constants can be found in Figure 2 (b).

		PSNR	MSSIM
Shape1	TVH1	30.9323	0.9812
	L0MS	31.7962	0.9837
	HOTVL	18.9389	0.9374
	ETV	36.2657	0.9967
	OUR	**36.8019**	**0.9973**
Shape2	TVH1	31.5509	0.9514
	L0MS	31.6024	0.9463
	HoTVL1	15.6438	0.8137
	ETV	36.6105	0.9972
	OUR	**37.6838**	**0.9975**

TABLE 1. PSNR and MSSIM of the reconstructed synthetic images.



(a) Shape1 (b) Shape2

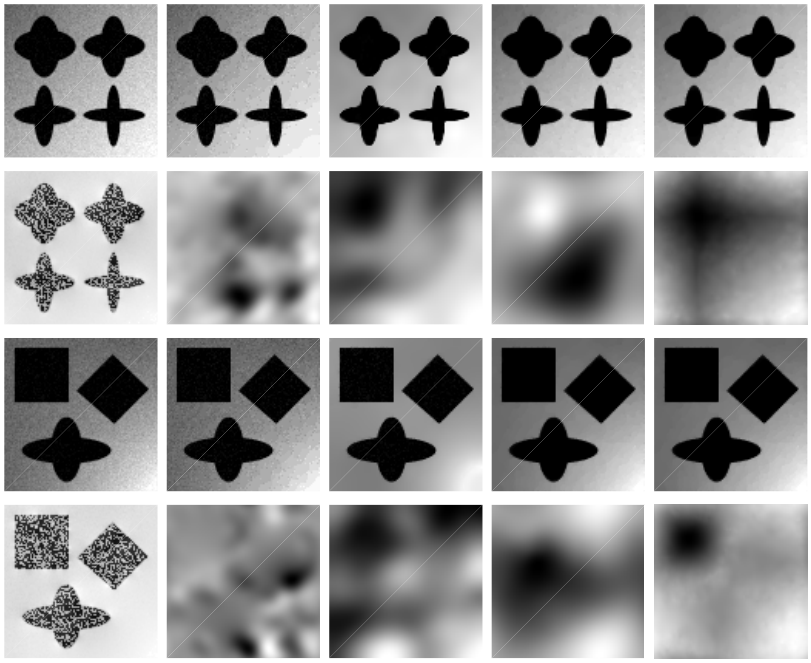
FIGURE 1. Two synthetic testing images.

4.1.1. ***Various Gaussian noises.*** In the medical imaging, the obtained images may be corrupted by the bias field due to non-uniform illuminations or the parallel magnetic resonance imaging (MRI)[3]. The correction of the bias field is similar to the retinex problem where we need to remove the light effect caused by the illumination. Here, we use two slices of T_1-weighted brain MR image as examples, which are contaminated the white Gaussian noise with different standard deviation as 0.03, 0.05, 0.07, 0.09 and the intensity non-uniformity as 40%. The quantitative comparisons as shown in Table 2 indicate that our proposed model achieves the best results in terms of the PSNR and the MSSIM.

To show the effectiveness via the visual quality of the retinex images, Figure 3 shows the related comparisons between our proposed model and other models. It is easy to observe that all models can provide visual preferable results compared with the original images. Specially, removing the noise by using our proposed model and the ETV is more robust than other models. In order to show the robustness of our proposed model, we zoom a part of the retinex image degraded by the noise with the standard deviation as 0.09 as an example and then plot them in the bottom column of Figure 3. Again, we still observe that the ETV and our proposed model can reconstruct more reasonable images. The important difference is that our proposed model gives more natural structures such edges and the piecewise constant regions.

4.1.2. ***Intensity inhomogenous images.*** Here we select one slice from the T_1-weighted brain volume and then generate ten degraded images by adding the white Gaussian noise with the standard deviation $\sigma = 0.001$ and the different intensity inhomogeneities. The PSNR and the MSSIM for the reconstructed images by using

[3]The MRIs can be downloaded from the website: *http://www.bic.mni.mcgill.ca/brainweb/*.



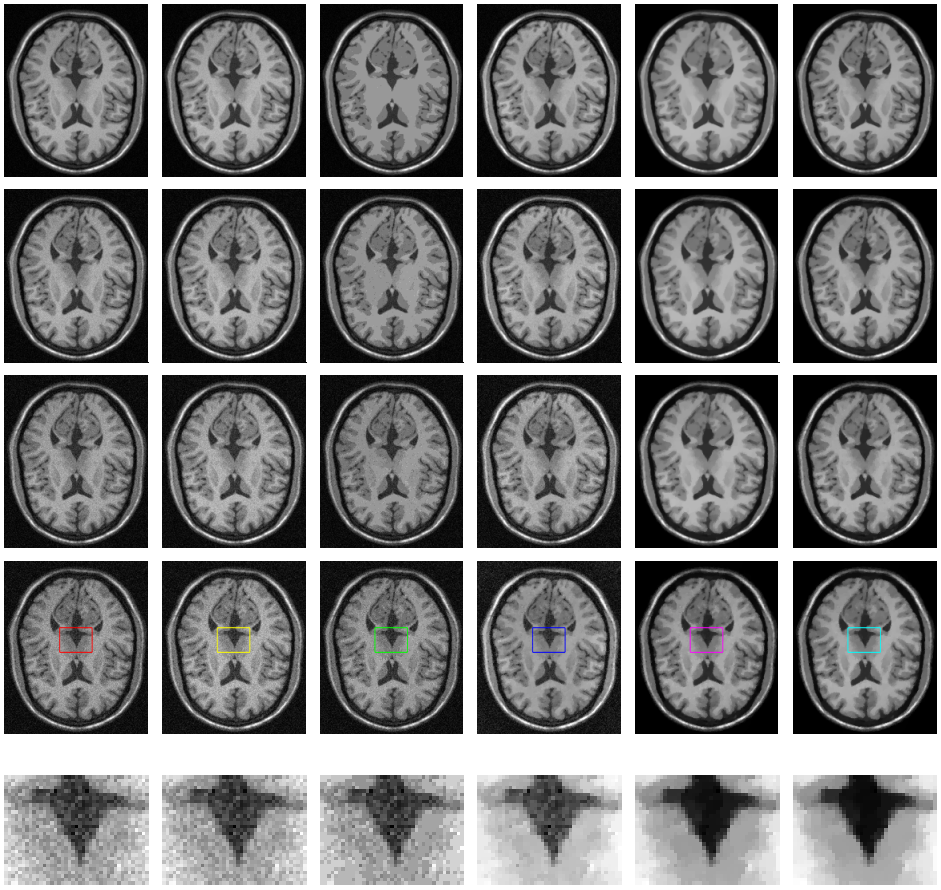
(a) TVH1 (b) L0MS (c) HOTVL1 (d) ETV (e) OUR

FIGURE 2. Reconstruction images based on the different model. The first and third rows: recovered reflectance r by five methods. The second and fourth rows: illumination l by five methods.

		0.03		0.05		0.07		0.09	
		PSNR	MSSIM	PSNR	MSSIM	PSNR	MSSIM	PSNR	MSSIM
Test Image1	TVH1	26.8509	0.9436	25.3513	0.9176	24.5925	0.8939	23.4850	0.8642
	HoTVL1	30.2055	0.9515	27.7438	0.9288	26.4877	0.9069	24.7440	0.8774
	L0MS	29.7718	0.9227	27.8711	0.9088	26.1986	0.8936	24.3809	0.8649
	ETV	32.6699	0.9904	31.1178	0.9844	29.1294	0.9767	28.4238	0.9686
	OUR	**33.0513**	**0.9909**	**31.2246**	**0.9846**	**30.1037**	**0.9781**	**28.4666**	**0.9686**
Test Image2	TVH1	27.4061	0.9230	26.5258	0.8935	25.5879	0.8661	24.0629	0.8348
	HoTVL1	30.0791	0.9317	29.0495	0.9065	27.0113	0.8795	25.1211	0.8481
	L0MS	32.0987	0.9246	29.0807	0.9016	27.1344	0.8760	24.9108	0.8382
	ETV	33.4363	0.9911	31.4108	0.9842	29.8199	0.9764	28.6496	0.9698
	OUR	**33.9694**	**0.9916**	**31.4609**	0.9841	**29.8766**	0.9760	**29.0947**	**0.9703**

TABLE 2. PSNR and MSSIM of T_1-weighted brain MRIs with different levels of Gaussian white noises.

different methods can be found in Figure 4. It obviously concludes that our proposed model outperforms other four models based on these quantitative criterions. Meanwhile, it can be observed that our proposed model is more stable than other models. That is to say the change of the intensity inhomogeneities can not seriously effect the reconstruction effectiveness of our proposed model due to the gentle change of the PSNR and the MSSIM as shown in Figure 4.



(a) Input (b) TVH1 (c) L0MS (d) HOTVL1 (e) ETV (f) OUR

FIGURE 3. Comparison of five methods on T_1-weighted brain MRIs with different levels of noises.

In general, the coefficient of variation (CV) is a standardized measure of the dispersion of the probability or frequency distribution. Then the CV can be used to estimate the effectiveness of the different method for reconstructing the degraded image. Formally, the CV of each tissue T defined by

$$\mathrm{CV(T)} = \frac{\sigma(\mathrm{T})}{\mu(\mathrm{T})} \times 100\%,$$

where $\sigma(\mathrm{T})$ and $\mu(\mathrm{T})$ are the standard deviation and the mean of the intensities in the tissue T. Here we employ it for evaluating the White Matter (WM), the Gray Matter (GM) and the Cerebrospinal Fluid (CSF) on the bias corrected images. The related boxplots are shown in Figure 5. It is similarly observed that the CV values based on our proposed model outperform other models, which implies that our proposed model can effectively deal with noises and intensity inhomogeneities again.

Now we consider to compare the visualization through choosing three degraded images as shown in Figure 6. The reconstructed images and the different images between the clean images and the restored images are shown in Figure 7. We can

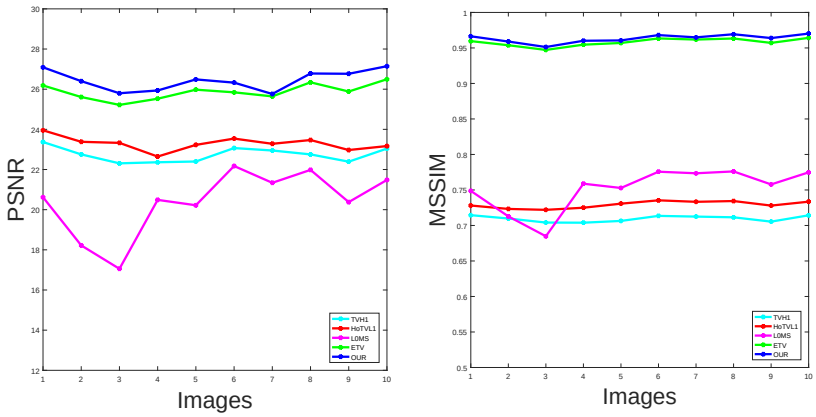


FIGURE 4. Comparison among five models in terms of the PSNR and the MSSIM. Here the x-axis denotes the the serial number of the image.

find that all methods can remove the noises and different intensity inhomogeneities efficiently. Especially, we can observe the colorbars of the different images based on our proposed method more dark and less structural information compared with other models, which implies the effectiveness of our proposed method. In addition, since the relative error between the clean images f and the restored images f_1

$$R(f, f_1) = \frac{\|f - f_1\|_2}{\|f\|_2}$$

can also display the effectiveness of the different model, we still use three images in Figure 6 as examples and then plot these curves of $R(f, f_1)$ as shown in Figure 8. It is obvious that $R(f, f_1)$ based on our proposed model is smaller than other methods. These illustrate the advantage and effectiveness of our proposed model again.

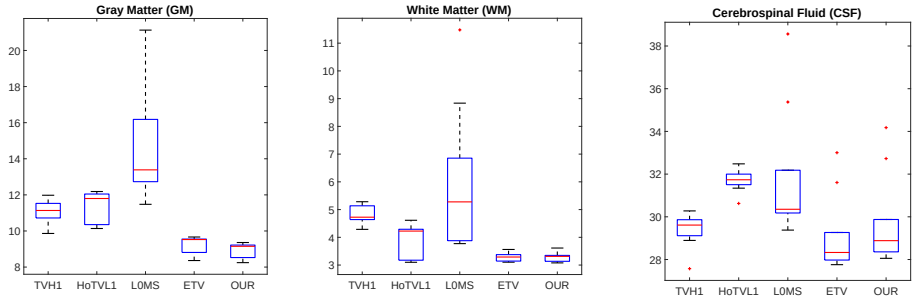
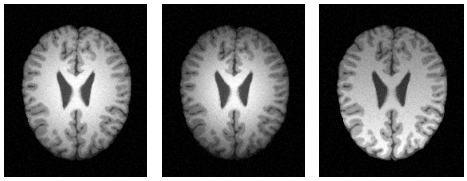


FIGURE 5. Comparison of the performance in terms of CV(%).

4.1.3. ***Convergence analysis***. Since our proposed model (8) is nonconvex and non-Lipschitz, it is not easy to prove the convergence of the proposed algorithm. In order to show the convergence from the numerical implementation, we need to



(a) Brain1 (b) Brain2 (c) Brain3

FIGURE 6. T_1-weighted brain images with different intensity inhomogeneities and noises.

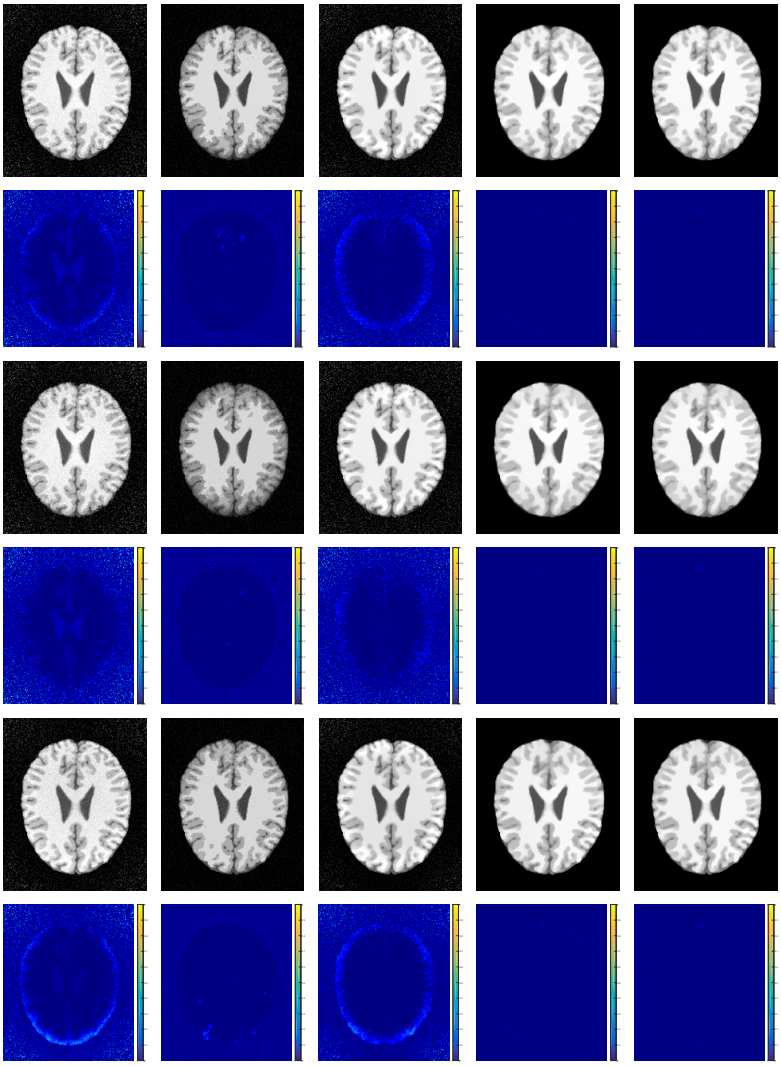
visually analyze the relative error of the solution r^n and l^n

$$R(r^n, l^n) = \Big(\frac{\|r^n - r^{n-1}\|_2^2}{\|r^n\|_2^2}, \frac{\|l^n - l^{n-1}\|_2^2}{\|l^n\|_2^2}\Big)$$

and the energy of the objective function in the problem (8). In general, $R(r^n, l^n)$ tending to zero implies the convergence of the numerical algorithm. Here we only use the third image as shown in Figure 6 as an example. As we can see from Figure 9 that $R(r^n, l^n)$ and the numerical energy decay when increasing the iteration, which demonstrates that our proposed algorithm converges well numerically.

4.2. **Real MRIs.** In this experiment, we present to reconstruct some real MRIs such as the bladder and brains as shown in Figure 10, and the heads as shown in Figure 13, where these real MRIs are obviously corrupted by the noise and the bias field. The reconstructed results presented in Figure 11 demonstrate that our proposed model can well recover the real intensities and improve the quality of the original MRIs. In addition, in order to show the reconstruction effectiveness, we select a slice of each image which locations are shown in Figure 10, and then plot their gray value curves as shown in Figure 12. Based on Figure 12(a), we can observe that the TVH1 and the HOTVL1 can not remove the bias field very well and the L0MS model makes the image too smooth. In contrast, our proposed model and the ETV can remove bias field and noises effectively. Furthermore, by the careful observation, the distribution of the gray values based on our proposed model is relatively closer to the original image than the ETV. That is to say, the reconstructed results generated by the ETV reduce the contrast of the images. In other words, our proposed model outperforms other models. Meanwhile, we can get similar conclusion from Figure 12(b) and (c).

Compared with the MRIs as shown in Figure 10, the MRIs of the head as shown in Figure 13 have more structure information. By observing the contour images of the original MRIs from the bottom of Figure 13, it is not easy to find the detailed information due to some effects of the bias field and the noise. These effects can reduce the effectiveness of the subsequent image processing such as the image segmentation and the image identification. Hence, it is very important to correct the bias field and the noise. In the following, we consider to reconstruct these MRIs. The reconstructed images and their contour images are respectively shown in Figure 14 and Figure 15. Figure 14 shows that the reconstructed comparisons generated by the different models can present visually preferable results compared with the original images. To intuitively observe the reconstructed results, we display the contours of the reconstructed results as shown in Figure 15. It is easy to observe some facts as follows. The TVH1 is able to remove non-uniform illuminations effectively, but it

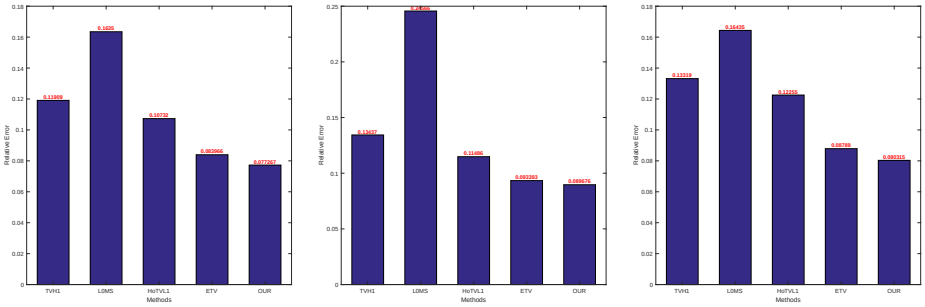


(a) TVH1 (b) L0MS (c) HOTVL1 (d) ETV (e) OUR

FIGURE 7. Performances of five methods on T_1-weighted brain images with different intensity inhomogeneities for the first, third and fifth rows; Colorbar to different between the clean images and restored images for the second, fourth and sixth rows.

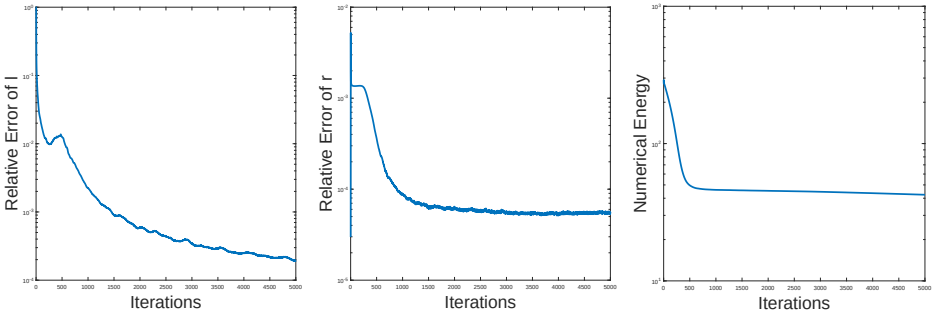
can break the background. The L0MS cannot recover the bottom of the corrupted images well. Compared with the HOTVL1 and the ETV, our proposed model not only removes the bias field more robust, but also preserves the details of the images.

4.3. **Illusion images.** To further demonstrate the superiority of our proposed model, we carry out the experiment by using two typical test images such as the Adelson's checkerboard shadow image and the Logvinenko's cube shadow image as examples. The degraded images are added to the white Gaussian noise with the



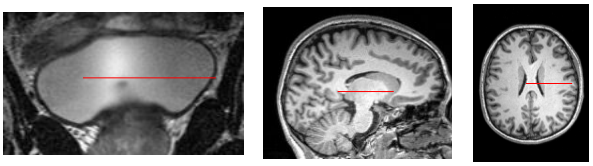
(a) (b) (c)

FIGURE 8. The relative errors of the original images and the restored images in Figure 6.(a) First image, (b) Second image and (c) Third image



(a) (b) (c)

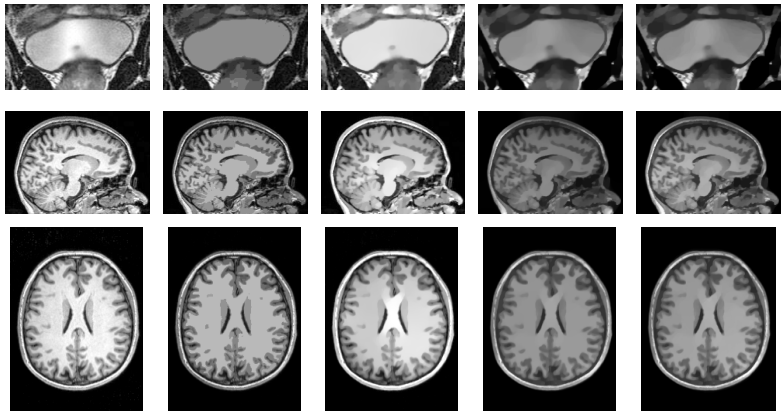
FIGURE 9. The relative errors of r and l and numerical energy of our model for the third image in Figure 6.



(a) Bladder (b) Brain1 (c) Brain2

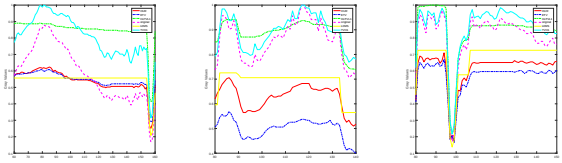
FIGURE 10. MRIs with noises and bias field.

standard deviation $\sigma = 0.001$ as shown in Figure 16. For the Adelson's checkerboard shadow image, though visually region B is brighter than region A, they are of the same intensity value. From the comparative results in Table 3, it can be observed that all of the PSNR and the MSSIM of the proposed model are higher than other models. In addition, the reflectance component R and the illumination component L obtained by different models are displayed in Figure 17. The visual results demonstrate that our proposed model can well decompose the reflectance and the illumination.



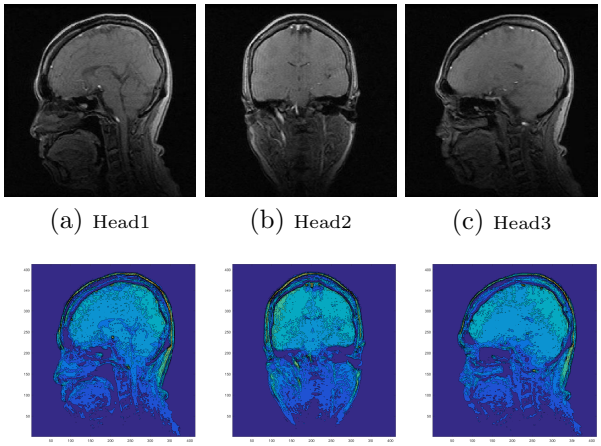
(a) TVH1 (b) L0MS (c) HOTVL1 (d) ETV (e) OUR

FIGURE 11. Bias field correction for the different MRIs.



(a) Bladder (b) Brain1 (c) Brain2

FIGURE 12. The gray values of underlined part with five methods.

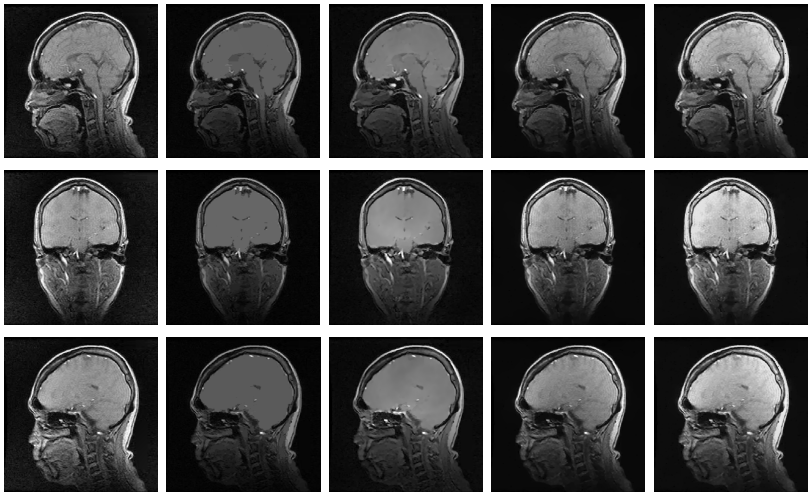


(a) Head1 (b) Head2 (c) Head3

(d) Contour1 (e) Contour2 (f) Contour3

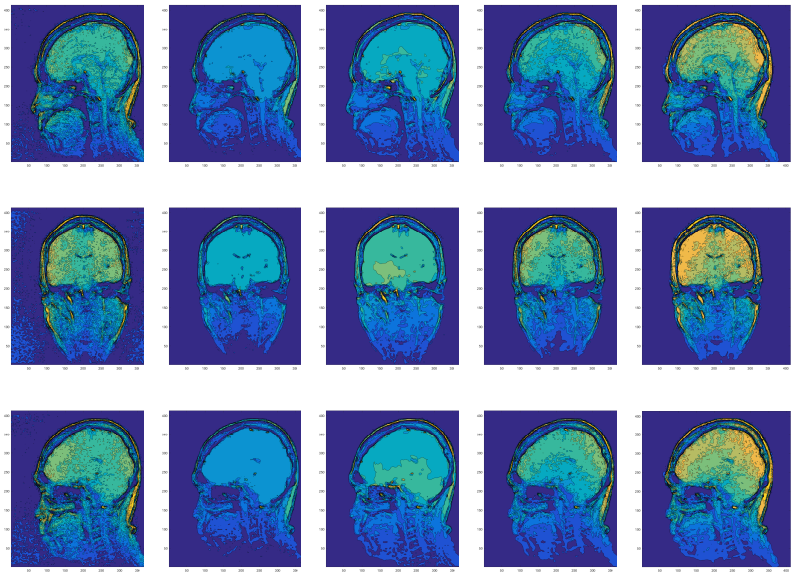
FIGURE 13. Original corrupted images and the corresponding contour images.

5. **Conclusion.** In this paper, we presented an efficient model to decompose the reflectance and the illumination based on the TV_p-HOTV_q regularization. Since the proposed model is non-convex, non-smoothing and non-Lipschitz, we employ the iteratively reweighed ℓ_1 algorithm based on the alternating minimization algorithm



(a) TVH1 (b) L0MS (c) HOTVL1 (d) ETV (e) OUR

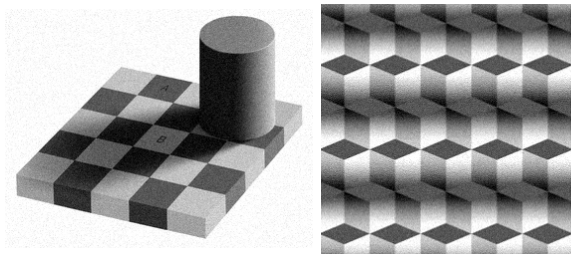
FIGURE 14. Comparisons of restored images with five methods. The head 1 for the first row, the head 2 for the second row and the head 3 for the third row.



(a) TVH1 (b) L0MS (c) HOTVL1 (d) ETV (e) OUR

FIGURE 15. FIGURE 14 corresponding contour images. The head 1 for the first row, the head 2 for the second row and the head 3 for the third row.

to solve it. Some mathematics properties of the proposed model and the numerical method were discussed. Various numerical results were implemented to demonstrate the advantages of our proposed model over some state-of-the-art retinex methods.



(a) Checkerboard (b) Cube

FIGURE 16. Two test images for the illusion problem. (a) Adelsons checkerboard shadow image. (b) Logvinenkos cube shadow image.

		PSNR	MSSIM
Checkboard	TVH1	31.7514	0.8887
	L0MS	33.1312	0.9602
	HOTVL1	29.5090	0.8725
	ETV	38.7053	0.9936
	OUR	**39.2105**	0.9934
Cube	TVH1	30.7779	0.9799
	L0MS	30.0935	0.9774
	HOTVL1	27.9950	0.9692
	ETV	29.6898	0.9866
	OUR	**33.3660**	**0.9896**

TABLE 3. PSNR and MSSIM of Retinex illusion images.

REFERENCES

[1] R. Amestoy, E. Provenzi, M. Bertalmío, and V. Caselles. *A perceptually inspired variational framework for color enhancement.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 31(3):458-474, 2009.

[2] M. Benning, F. Knoll, C. Schonlieb, and T. Valkonen. *Preconditioned ADMM with nonlinear operator constraint.* IFIP Conference on System Modeling and Optimization, 494:117-126, 2015.

[3] M. Bertalmío, V. Caselles, E. Provenzi, and A. Rizzi. *Perceptual color correction through variational techniques.* IEEE Transactions on Image Processing, 16(4):1058-1072, 2007.

[4] M. Bertalmío, V. Caselles, and E. Provenzi. *Issues about retinex theory and contrast enhancement.* International Journal of Computer Vision, 83(1):101-119, 2009.

[5] A. Blake. *Boundary conditions for lightness computation in Mondrian world.* Computer Vision Graphics Image Processing, 32(3):314-327, 1985.

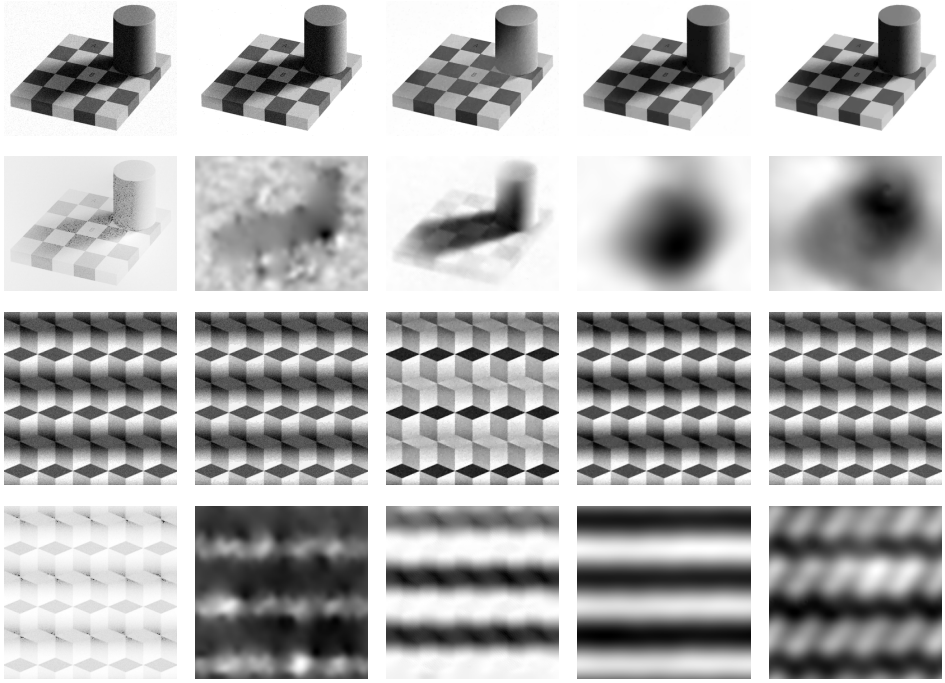
[6] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. *Distributed optimization and statistical learning via the alternating direction method of multipliers.* Foundations and Trends in Machine Learning, 3(1):1-122, 2011.

[7] E. Candes, M. Wakin, and S. Boyd. *Enhancing sparsity by reweighted l_1 minimization.* Journal of Fourier Analysis and Applications, 14(5):877-905, 2008.

[8] T. Cooper and F. Baqai. *Analysis and extensions of the frankle-mccann retinex algorithm.* Journal of Electronic Imaging, 13(1):85-93, 2004.

[9] X. Chen, F. Xu, and Y. Ye. *Lower bound theory of nonzero entries in solutions of $\ell^2 - \ell^p$ minimization.* SIAM Journal on Scientific Computing, 32(5):2832-2852, 2010.

[10] X. Chen and W. Zhou. *Convergence of the reweighted ℓ_1 minimization algorithm for $\ell_2 - \ell^p$ minimization.* Journal Computational Optimization and Application, 59(1-2):47-61, 2014.



(a) TVH1 (b) L0MS (c) HOTVL1 (d) ETV (f) OUR

FIGURE 17. Decomposition comparison of the checkerboard image and the cube image. The first and third rows: recovered reflectance r by five methods. The second and fourth rows: illumination l by five methods.

[11] J. Douglas and H. Rachford. *On the numerical solution of heat conduction problems in two and three space variables.* Transactions of the American Mathematical Society, 421-439, 1976.

[12] Y. Duan, H. Chang, W. Huang, J. Zhou, Z. Lu, and C. Wu. *The L_0 regularized Mumford-Shah model for bias correction and segmentation of medical images.* IEEE Transactions on Image Processing, 24(11):3927-3938, 2015.

[13] D. Gabay and B. Mercier. *A dual algorithm for the solution of nonlinear variational problems via finite element approximations.* Computers and Mathematics with Applications, 2(1):17-40, 1976.

[14] N. Galatsanos and A. Katsaggelos. *Methods for choosing the regularization parameter and estimating the noise variance in image restoration and their relation.* IEEE Transactions on Image Processing, 1(3):322-336, 1992.

[15] D. Ghilli and K. Kunisch. *On monotone and primal-dual active set schemes for ℓ^p-type problems, $p \in (0, 1]$.* Computational Optimizationand Applications, 72(1):45-85, 2019.

[16] R. Glowinski, S. Luo, and X. Tai. *Fast operator-splitting algorithms for variational imaging models: Some recent developments.* Handbook of Numerical Analysis, 20:191-232, 2019.

[17] R. Glowinski, S. Osher, and W. Yin. *Splitting Methods in Communication, Imaging, Science, and Engineering.* Springer 2017.

[18] Z. Gu, F. Li, and X. Lv. *A detail preserving variational model for image Retinex.* Applied Mathematical Modelling, 68:643-661, 2019.

[19] P. Hansen and D. Leary. *The use of the L-curve in the regularization of discrete ill-posed problems.* SIAM Journal on Scientific Computing, 14(6):1487-1503, 1993.

[20] B. Horn. *Determining lightness from an image.* Computer Graphics Image Processing, 3(4):277-299, 1974.

[21] B. Horn. *Understanding image intensities.* Artificial intelligence, 8(2):201-231, 1977.

[22] D. Jobson, Z. Rahman, and G. Woodell. *Properties and performance of a center/surround retinex.* IEEE Transactions on Image Processing, 6(3):451-462, 1997.

[23] D. Jobson, Z. Rahman, and G. Woodell. *A multiscale Retinex for bridging the gap between color image and the human observation of scenes.* IEEE Transactions on Image Processing, 6(7):965-976, 2002.

[24] Y. Jung, T. Jeong, and S. Yun. *Non-convex TV denoising corrupted by impulse noise.* Inverse Problems and Imaging, 11(4):689-702, 2017.

[25] R. Kimmel, M. Elad, D. Shaked, R. Keshet, and I. Sobel. *A variational framework for retinex.* International Journal of Computer Vision, 52(1):7-23, 2003.

[26] M. Lai, Y. Xu, and W. Yin. *Improved iteratively reweighted least squares for unconstrained smoothed ℓ_q minimization.* SIAM Journal on Numerical Analysis, 51(2):927-957, 2013.

[27] E. Land and J. Mccann. *Lightness and Retinex theory.* Journal of the Optical Society of America, 61(1):1-11, 1971.

[28] E. Land. *The Retinex theory of color vision.* Scientific American, 237(6):108-128, 1977.

[29] E. Land. *An alternative technique for the computation of the designator in the retinex theory of color vision.* Proceedings of the National Academy of Sciences, 83(10):3078-3080, 1986.

[30] M. Langer and S. Zucker. *Spatially varying illumination: A computational model of converging and diverging sources.* European Conference on Computer Vision, 226-232, 1994.

[31] A. Lanza1, S. Morigi1, and F. Sgallari. *Constrained $TV_p - \ell_2$ model for image restoration.* Journal of Scientific Computing, 68(1):64-91, 2016.

[32] L. Liu, Z. Pang, and Y. Duan. *Retinex based on exponent-type total variation scheme.* Inverse Problems and Imaging, 12(5):1199-1217, 2018.

[33] Z. Liu, C. Wu, and Y. Zhao. *A new globally convergent algorithm for non-Lipschitz $\ell^p - \ell^q$ minimization.* Advances in Computational Mathematics, 45(3):1369-1399, 2019.

[34] J. Liang and X. Zhang. *Retinex by higher order total variation L^1 decomposition.* Journal of Mathematical Imaging and Vision, 52(3):345-355, 2015.

[35] Z. Lu. *Iterative reweighted minimization methods for ℓ^p regularized unconstrained nonlinear programming.* Mathematical Programming: Series A and B, 147(1-2):277-307, 2014.

[36] D. Marini and A. Rizzi. *A computational approach to color adaptation effects.* Image and Vision Computing, 18(13):1005-1014, 2000.

[37] W. Ma and S. Osher. *A TV bregman iterative model of retinex theory.* Inverse Problems and Imaging, 6(4):697-708, 2012.

[38] J. Mccann. *Lessons learned from mondrians applied to real images and color gamuts.* Proceedings of the IST/SID 7th Color Imaging Conference, 1-8, 1999.

[39] J. Morel, A. Petro, and C. Sbert. *Fast implementation of color constancy algorithms.* Proceedings of SPIE, 7241, 2009.

[40] J. Morel, A. Petro, and C. Sbert. *A PDE formalization of retinex theory.* IEEE Transactions on Image Processing, 19(11):2825-2837, 2010.

[41] V. Morozov. *Methods for Solving Incorrectly Posed Problems.* Springer-Verlag, New York, 1984.

[42] M. Ng and W. Wang. *A total variation model for retinex.* SIAM Journal on Imaging Sciences, 4(1):345-365, 2011.

[43] P. Ochs, A. Dosovitskiy, T. Brox, and T. Pock. *On iteratively reweighted algorithms for nonsmooth nonconvex optimization in computer vision.* SIAM Journal on Imaging Sciences, 8(1):331-372, 2015.

[44] J. Oliveira, J. Dias, and M. Figueiredo. *Adaptive total variation image deblurring: A majorizationCminimization approach.* Signal Processing, 89(9):1683-1693, 2009.

[45] H. Pan, Y. Wen, and H. Zhu. *A regularization parameter selection model for total variation based image noise removal.* Applied Mathematical Modelling, 68:353-367, 2019.

[46] E. Provenzi, D. Marini, L. De Carli, and A. Rizzi. *Mathematical definition and analysis of the retinex algorithm.* Journal of the Optical Society of America A, 22(12):2613-2621, 2005.

[47] S. Sabacha and M. Teboulle. *Lagrangian methods for composite optimization.* Handbook of Numerical Analysis, 20:401-436, 2019.

[48] A. Theljani and K. Chen. *A Nash game based variational model for joint image intensity correction and registration to deal with varying illumination.* Inverse Problems, 36(3):034002, 2020.

[49] Y. Wen and R. Chan. *Using generalized cross validation to select regularization parameter for total variation regularization problems.* Inverse Problems and Imaging, 12(5):1103-1120, 2018.

[50] W. Wang and C. He. *A variational model with barrier functionals for Retinex.* SIAM Journal on Imaging Sciences, 8(3):1955-1980, 2015.

[51] J. Zhang, R. Chen, C. Deng, and S. Wang. *Fast linearized augmented method for Euler's elastica model.* Numerical Mathematics:Theory, Methods and Applications, 10(1):98-115, 2017.

[52] X. Zhang, Y. Shi, Z. Pang, and Y. Zhu. *Fast algorithm for image denoising with different boundary conditions.* Journal of the Franklin Institute, 354:4595-4614, 2017.

[53] D. Zosso, G. Tran, and S. Osher. *A unifying retinex model based on non-local differential operators.* Computational Imaging XI, 865702, 2013..

[54] W. Zuo, D. Meng, L. Zhang, X. Feng, and D. Zhang. *A generalized iterated shrinkage algorithm for non-convex sparse coding.* IEEE International Conference on Computer Vision, 217-224, 2013.

Received October 2019;