

Image Enhancement and Restoration: State of the Art of Variational Retinex Models

Guojia Hou*, Guodong Wang, Zhenkuan Pan, Baoxiang Huang, Huan Yang and Teng Yu

Abstract—The image we interested on may suffer the problems of low contrast, non-uniform lighting, blurring and so on, which will affect its interpretation and recognition. Therefore, enhancing and restoring such a low-quality image is of significance. Retinex models are powerful methods for image enhancement and restoration. In recent years, variational methods based on Retinex theory are increasingly paid more attention. In this paper, we first summarize a short review of traditional Retinex methods and then present existing variational Retinex models. We also present detailed discretisation of these models and numerical implementation of the split Bregman algorithm for solving these models using the fast Fourier transform. We further demonstrate the advantages and disadvantages of several representative variational models in the context of image enhancement through extensive experiments. These models and techniques can also be used for other applications, such as image decomposition, inpainting and segmentation.

Index Terms—Image enhancement, restoration, retinex, variational model, split bregman

I. INTRODUCTION

IMAGE enhancement and restoration are essential preprocessing parts of many image analysis processes such as image segmentation [1], [2], image recognition [3], [4], and so on. To address the difficulties of image enhancement and restoration, a lot of methods have been proposed. Some methods [5]-[7] manipulate image histogram to enhance images, some methods apply a nonlinear transformation [8] to compress dynamic range of an image, and some methods [9] develop a gray level transformation for an image. Among these color image enhancement and restoration approaches, technologies based on the Retinex theory have been widely used in many fields. The Retinex theory is originally proposed by Land and McCann [10] as a color perception model of the human visual system (HVS). Our visual system ensures that the perceived color of one object remains relatively constant

Manuscript received January 4, 2017; revised May 2, 2017. The research work is supported by China Postdoctoral Science Foundation (No. 2017M612204) and the National science and technology support project of China (2014BAG03B05), and in part by the National Natural Science Foundation of China (No.61602269).

Guojia Hou is with College of Computer Science & Technology and College of Automation and Electrical Engineering, Qingdao University, China (corresponding author, e-mail:hgjouc@126.com).

Guodong Wang, Zhenkuan Pan, Baoxiang Huang and Huan Yang are with College of Computer Science & Technology, Qingdao University, China.

Teng Yu is with School of Electronic and Information Engineering, Qingdao University, China.

with different illuminations conditions, which called color constancy. The fundamental assumption of Retinex theory is that a given image S can be decomposed into two different images: the reflectance R and the illumination L such that

$$S(x) = R(x) \cdot L(x) \quad (1)$$

at each pixel (x, y) in the image domain. To simplify equation (1), we can apply the logarithm operation on R and L which leads to

$$s(x) = r(x) + l(x) \quad (2)$$

where $s = \log S$, $r = \log R$, and $l = \log L$.

Fig. 1(a) presents the well-known Retinex illusions called “Adelson’s checker shadow illusion”. For us, region A seems dark than region B, but actually the intensity value of them are exactly equal as shown in Fig. 1(b). This is because they are in different illumination conditions. It is noted that region B is under the shadow of a green cylinder so that the illumination in the region B is weaker than that in the region A, i.e., $L(B) < L(A)$. Based on the Retinex theory, the reflectance in the region B is larger than that in the region A, i.e., $R(A) < R(B)$ to ensure $S(A) = S(B)$. As only the reflectance R is perceived by the HVS, that is why the region A seems darker for us.

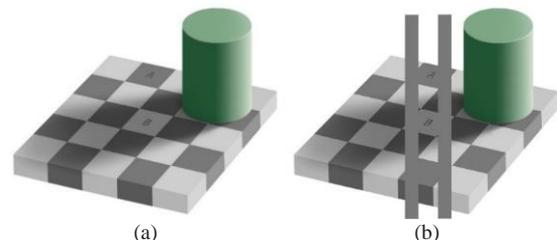


Fig. 1. Adelson’s checker shadow illusion: (a) Original image, (b) Demonstration.

To simulate the mechanism of HVS, it needs to recover the reflectance R from a given image S , which is mathematically an ill-posed problem. Fig. 2 illustrates the general process of Retinex algorithms. Base on Land and McCann’s Retinex, a number of algorithms are proposed in the literature for its solution vary in their way of overcoming this limitation. These approaches can be classified into five categories such as path-based algorithms [10]-[12], recursive algorithms [13]-[15], center/surround algorithms [16]-[18], PDE-based algorithms [19]-[21] and variational algorithms [22]-[24].

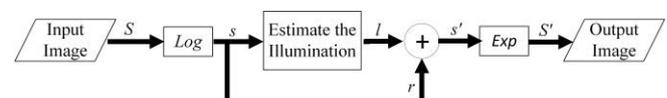


Fig. 2. The general flow chart of Retinex algorithms.

The original Retinex algorithm is a typical path based algorithm, which performs a random choice on both starting pixels and path routes. The main problems of this kind of algorithms are their high computational complexity and free parameters. Afterwards, Frankle and McCann [13], [14] applies recursive matrix calculation to replace path computation to boost its efficiency. However, the disadvantage of these algorithms is that their computation results are strongly relied on the iteration number. As well known, the single scale Retinex (SSR) [16] and the multiscale Retinex (MSR) [17] are the representative algorithms of the center/surround approaches. As their simplicity and effectiveness, many improvement algorithms [25]-[30] are proposed on the basis of SSR and MSR. For instance, the multiscale Retinex with color restoration (MSRCR) [18] method is an improvement of the MSR while it does not solve the problem of the halo artifacts near strong edges. In PDE-based algorithm [19], it assumes that the reflectance changes at sharp edges and the illumination varies smoothly, therefore the above ill-posed problem can be modeled as a Poisson equation. On the basis of the same assumptions in [19], variational Retinex models [22], [24] are established by employing various regularization in terms of illumination and reflection functions. After that, many variational algorithms are further proposed on the basis of this variational framework. In this paper, we mainly focus on the state-of-the-art variational Retinex models.

The outline of this paper is as follows: we start by presenting a short review of some traditional Retinex flavors in section 2. Section 3 summarizes many variational Retinex models and numerical algorithms. Section 4 briefly introduces some recent nonlocal variational Retinex models. Experimental results and comparisons are shown in section 5. Finally, we conclude this paper in section 6.

II. TRADITIONAL RETINEX ALGORITHMS

Here, the traditional Retinex algorithms are basically categorized as three types: path-based algorithms, center/surround algorithms and PDE-based algorithms. In this section, we will provide a brief introduction of these algorithms.

A. Path-based Algorithms

Path-based algorithms are the primary methods in Retinex theory, which have high computation complexity and need too many parameters. The difference among them is the way and the order to choose the other comparison pixels.

Land's original scheme [10] applied stochastic theory considering the reflectance at each pixel relying on the multiplication of the ratios along random walks. Further developments introduced random distributed path algorithm with Brownian motion. Brainard and Wandell [31] studied the convergence properties for a large set of long paths using stochastic theory and found that the output would be a degraded version of the given image. Firstly, a large number of walks are initiated at random locations of an input image s ($\log S$), adopting the gray value of their initial position. Then, an accumulator image s' that has the same size as s is initialized to zero. As the walkers walked around, the image s' is updated by adding their values to each position they visited.

Finally, the reflectance image is obtained by normalizing the accumulator image.

Frankle and McCann [13], [15] extended the path-based algorithms and employed an efficient recursive matrix calculation to replace the path computation. In their algorithms, long distance interactions are first calculated and then progressively moves to short-distance interactions with decreasing the spacing. At each step, the illumination estimate is updated by resetting a ratio product. The illumination image \hat{l}_0 is initialized to be s , the original image. The algorithm performs the following iterative procedure,

$$\hat{l}_{n+1} = \max \left\{ \frac{\hat{l}_n + s}{2}, \frac{\hat{l}_n + D_n[\hat{l}_n]}{2} \right\} \quad (3)$$

where D_n is a translation operator, shifting the image by the n th element of a sequence of spirally decaying translation vectors. The size of the first displacement is set to be half the minimum between the image width and height.

More recent variants of the algorithms involving multi-resolution image pyramids [13], [14], different sampling patterns [32], [33], or ratio modifiers [34] are proposed. Provenzi et al. [35] replaced the path-based sampling pattern with a repeated sampling through random sprays and formalized the standard path-based sampling process representation of the Retinex model. It proves that despite the overall path dependence, this model can be given a representation in terms of absorbing Markov chains, by means of the embedding into a suitable state space. G. Gianini et al. [36] derived the corresponding analytic model, accounting for the combined effects of path function, path sampling process and starting point sampling process. Here, a numerical algorithm named ReMark (reset only Retinex by absorbing Markov chains) was provided for computing this new representation.

B. Center/Surround Algorithms

The center/surround approach was proposed by Land, et al. [10] and later improved by Jobson, et al. [37]. Jobson et al. then proposed the SSR (single scale Retinex) and the MSR (multiscale Retinex) which are on the basis of the assumption that the illumination component tends to vary smoothly, while the reflectance changes at sharp edges. Therefore the output reflectance values can be obtained by subtracting a blurred given image. These algorithms are easily implemented but need a large number of parameters. The SSR algorithm is given by

$$R_{SSR}(i) = \log I(i) - \log [F * I](i) \quad (4)$$

where $R(i)$ is the Retinex output, $I(i)$ is the image distribution in the i th spectral band, and F is a Gaussian kernel. The MSR algorithm is simply the combination of different SSRs:

$$R_{MSR}(i) = \sum_{n=1}^n w_n R_n(i) = \log I(i) - \sum_{n=1}^n w_n \log [F_n * I](i) \quad (5)$$

where $w_n : \sum_{n=1}^n w_n = 1$ are the weights of each scale and F_n are Gaussian kernels of different scales.

By changing the order of log and Gaussian convolution of the SSR algorithm (4), one gets homomorphic filtering (HF)

$$R_{HF}(i) = \log I(i) - [F * \log I](i) \quad (6)$$

Given the kernel $w(i, j) \equiv F$ and $f \equiv \log$, it can be identified as

$$\begin{aligned} R(i) &= \sum_{j=1}^j w(i, j) \log \left(\frac{I(i)}{I(j)} \right) \\ &= \log I(i) - \sum_{j=1}^j w(i, j) \log I(j) \end{aligned} \quad (7)$$

C. PDE-based Algorithms

In PDE-based formulations, people usually take a logarithm for the original formulation as $s = i + r$ where $s = \log(s)$, $r = \log(R)$, $i = \log(I)$. As the illumination i of an input image is supposed to be changing smoothly, the spatial derivatives of the observed intensity are primarily affected the reflectance r . Horn [21] applied the Laplacian to obtain $\Delta s = \Delta i + \Delta r$: Δi will be finite anywhere, while Δr will be zero excepting edge regions. So Laplacian is supposed to yield the reflectance without regard to the finite parts of the observed intensity.

$$\Delta r = \delta_r(\Delta s) \quad (8)$$

where the threshold function is

$$\delta_r(x) = \begin{cases} x & \text{if } |x| > t \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

Horn's algorithm has been strongly backed up by Morel, Petro and Sbert [38], [39], which shows a very tight connection with Land's original, reset free Retinex, and Horn's Laplacian thresholding algorithm. Actually, their comment actually confirms the mathematical result: under the Brownian path assumption the original Retinex becomes a Poisson equation, very similar to Horn's physical model and almost identical to Blake's model [20].

In a word, in the PDE-based formulations [20], [21], [38], [39], threshold functions are mainly applied on eliminating the effect of illumination. Then the reflectance can be recovered with Poisson equations which can be solved by several effective algorithms such as fast Fourier transform (FFT).

III. VARIATIONAL RETINEX MODELS

In recent two decades, numerous variational Retinex models have been applied to image enhancement and restoration. The primary variational model for the Retinex problem was proposed by Kimmel et.al [22] in 2003. Subsequently, variations Retinex models have been proposed based on this variational theory. First, Ma and Osher [40] established a total variation (TV) by dropping a few terms and replacing Gaussian smoothness of the reflectance. As a complication, they also introduced a nonlocal TV regularizer to replace the local TV prior. Afterwards, Ng and Wang [41] presented an improved total variation model for Retinex by a L^2 fidelity prior between reflectance and intensity. Chen et al. proposed a new logarithmic total variation (LTV) model based on minimizing an L^1 norm. They also modified the L^1 Retinex model for recovering magnetic resonance imaging

(MRI) [42]. In this section, we summarize several variational Retinex models and numerical algorithms.

A. Kimmel's Variational Model (H^1 - L^2)

Similar to traditional Retinex algorithms, Kimmel's variational model assumes spatial smoothness of the illumination field. In addition, knowledge of the limited dynamic range of the reflectance is used as a constraint in the recovery process. Three penalty terms are included, the first one forces spatial smoothness on the illumination, the second focus a proximity between l and s , and the third adopts a Bayesian view point of the estimation problem, which leads to an algebraic regularization term that contributes to better conditioning of the reconstruction problem.

The penalty functional can be expressed as following according to all the above assumptions.

$$\hat{l} = \arg \min_l \int_{\Omega} (|\nabla l|^2 + \alpha(l-s)^2 + \beta|\nabla(l-s)|^2) dx \quad (10)$$

subject to (s.t.) $l > s$, and $\langle \nabla l, \mathbf{n} \rangle = 0$ on $\partial\Omega$,

where $\Omega \rightarrow \mathbb{R}$ is the support of the image, $\partial\Omega$ is its boundary, and \mathbf{n} is the normal to the boundary. α and β are free nonnegative real parameters.

This is a quadratic programming problem that can be solved by many methods such project normalized steepest descent method as in paper [22].

B. Total Variation (TV) Model

In 2010, Ma and Osher [40] reduced the penalty terms and replaced H^1 -smoothness of the reflectance using a TV regularized model.

$$\hat{r} = \arg \min_r \int_{\Omega} \left(t|\nabla r| + \frac{1}{2}|\nabla r - \nabla s|^2 \right) dx \quad (11)$$

which was implemented by split Bregman iterations.

$$\hat{r} = \arg \min_r \int_{\Omega} \left(t\sqrt{d_x^2 + d_y^2} + \frac{1}{2}\|\mathbf{d} - \nabla s\|_2^2 \right) \quad \text{s.t. } \mathbf{d} = \nabla r \quad (12)$$

By the general split Bregman method, we have

$$\begin{aligned} \mathbf{d}^{k+1} &= \arg \min_{\mathbf{d}} \frac{t}{\lambda} \int_{\Omega} \sqrt{d_x^2 + d_y^2} + \frac{1}{2\lambda} \|\mathbf{d} - \nabla s\|_2^2 \\ &\quad - \langle \mathbf{b}^k, \mathbf{d} - \mathbf{d}^k \rangle + \frac{1}{2} \|\mathbf{d} - \nabla r^k\|_2^2 \end{aligned} \quad (13)$$

$$r^{k+1} = \arg \min_r \langle \mathbf{b}^k, \nabla(r - r^k) \rangle + \frac{1}{2} \|\mathbf{d}^{k+1} - \nabla r\|_2^2$$

$$\mathbf{b}^{k+1} = \mathbf{b}^k - (\mathbf{d}^{k+1} - \nabla r^{k+1})$$

The above equation (13) can be explicitly solved [43]. They are summarized in the following algorithms:

Initialize $r^0 = s, \mathbf{d}^0 = \mathbf{b}^0 = 0$

While $\frac{\|r^{k+1} - r^k\|_2}{\|r^{k+1}\|_2} > \varepsilon$

$$(1) \mathbf{d}^{k+1} \leftarrow \frac{1}{1+\lambda} \text{shrink}_t \left(\nabla s + \lambda \nabla u^k + \lambda \mathbf{b}^k \right) \quad (14)$$

$$(2) r^{k+1} \leftarrow \text{solution of } \Delta r = \left(\text{div}(\mathbf{d}^{k+1} - \mathbf{b}^k) \right)$$

$$(3) \mathbf{b}^{k+1} \leftarrow \mathbf{b}^k - (\mathbf{d}^{k+1} - \nabla u^{k+1})$$

end While

The isotropic shrinkage function is defined by

$$\text{shrink}_t(z) = \begin{cases} 0 & \text{if } \|z\|_2 \leq t \\ z - z \frac{t}{\|z\|_2} & \text{if } \|z\|_2 > t \end{cases} \quad (15)$$

C. TV-L² Model

Ng and Wang [41] presented an L²-fidelity prior between reflectance and intensity. Unlike Ma and Osher model, the reflection function is considered, and some constraints and a fidelity term are added in the proposed energy functional:

$$\begin{aligned} (\hat{r}, \hat{l}) &= \arg \min_{(r,l) \in \Lambda} E(r,l) \\ &= \int_{\Omega} |Dr| + \frac{\alpha}{2} \int_{\Omega} |\nabla l|^2 dx + \frac{\beta}{2} \int_{\Omega} (l-r-s)^2 dx + \frac{\mu}{2} \int_{\Omega} l^2 dx \end{aligned} \quad (16)$$

with $\Lambda = \{(r,l) \mid (r,l) \in BV(\Omega) \times W^{1,2}(\Omega), r \geq 0, l \geq s\}$

Here α , β and μ are positive numbers for regularization parameters, the term $\int_{\Omega} (l-r-s)^2 dx$ is utilized for the fidelity, and the term $\int_{\Omega} l^2 dx$ is used only for the theoretical setting.

A fast alternating minimization scheme is employed to solve the proposed model.

Step 1. Set $k=0$, and let $l^0 = s$ be the initial illumination function.

Step 2. At the k th iteration:

(1) Given l^k , compute $r^{k+\frac{1}{2}}$ by solving

$$\hat{r} = \arg \min_r E_1(r) = \int_{\Omega} |Dr| + \frac{\beta}{2} \int_{\Omega} (r+s-l^k)^2 dx \quad (17)$$

then update r^{k+1} by using

$$r^{k+1} = \max \left\{ r^{k+\frac{1}{2}}, 0 \right\} \quad (18)$$

(2) Given r^{k+1} , compute $l^{k+\frac{1}{2}}$ by solving

$$\begin{aligned} \hat{l} &= \arg \min_l E_2(l) = \frac{\alpha}{2} \int_{\Omega} |\nabla l|^2 dx \\ &+ \frac{\beta}{2} \int_{\Omega} (l-r^{k+1}-s)^2 dx + \frac{\mu}{2} \int_{\Omega} l^2 dx \end{aligned} \quad (19)$$

Step 3. Go back to step 2 until satisfy the following inequality.

$$\frac{\|l^{k+1} - l^k\|}{\|l^{k+1}\|} \leq \varepsilon_l, \quad \frac{\|r^{k+1} - r^k\|}{\|r^{k+1}\|} \leq \varepsilon_r \quad (20)$$

The minimization subproblem in equation (17) can be translated using the split Bregman method by introducing the auxiliary variable d in the calculation:

$$\min_{r,d} \left\{ \int_{\Omega} |d| + \frac{\beta}{2} \int_{\Omega} (r+s-l^k)^2 dx \right\} \text{ s.t. } d = \nabla r \quad (21)$$

The computation procedure of above equation is presented as follows.

Step 1. Let $w^0 = 0$, $\mathbf{b}^0 = (b_x^0, b_y^0) = 0$ be the initial value.

Step 2. At the i th iteration:

(1) Given w^i and \mathbf{b}^i , update \mathbf{d}^{i+1} by using

$$\mathbf{d}^{i+1} = \text{shrink} \left(\nabla w^i + \mathbf{b}^i, \frac{1}{\lambda} \right) \quad (22)$$

(2) Given \mathbf{d}^{i+1} and \mathbf{b}^i , update w^{i+1} by using

$$w^{i+1} = F^{-1} \left(\frac{\beta F(l^k - s) + \lambda (F * (\nabla_x) F(d_x^{i+1} - b_x^i) + F * (\nabla_y) F(d_y^{i+1} - b_y^i))}{\lambda (F * (\nabla_x) F(\nabla_x) + F * (\nabla_y) F(\nabla_y)) + \beta} \right) \quad (23)$$

where F is the discrete Fourier matrix.

(3) Given w^{i+1} and \mathbf{d}^{i+1} , update \mathbf{b}^{i+1} by using

$$\mathbf{b}^{i+1} = \mathbf{b}^i - (\mathbf{d}^{i+1} - \nabla w^{i+1}) \quad (24)$$

Step 3. Set $r^{k+\frac{1}{2}} = w^{i+1}$ until $\frac{\|w^{i+1} - w^i\|}{\|w^{i+1}\|} \leq \varepsilon_w$

The minimization subproblem in equation (19) can be solved by using fast Fourier transform (FFT):

$$l^{k+1} = F^{-1} \left(\frac{\beta F(r^{k+1} + s)}{\alpha (F * (\nabla_x) F(\nabla_x) + F * (\nabla_y) F(\nabla_y)) + \beta + \mu} \right) \quad (25)$$

D. L¹ Variational Model

In 2011, Ma and Osher [42] established an L¹ based variational Retinex model. In this paper, instead of minimizing the L² norm, its motivation is to minimize the L¹ norm of $\nabla r - \delta_t(\nabla s)$. Namely, the L¹ Retinex model is

$$\hat{r} = \arg \min_r \int_{\Omega} |\nabla r - \delta_t(\nabla s)| \quad (26)$$

δ_t is a threshold function applied on ∇r to preserve the gradient of reflectance, and it is defined as

$$\delta_t(\mathbf{z}) = (\tau_t(z_1), \dots, \tau_t(z_n)) \quad (27)$$

with

$$\tau_t(z) = \begin{cases} z & \text{if } |z| > t \\ 0 & \text{otherwise} \end{cases} \quad (28)$$

where the threshold t is a positive parameter. Therefore, given a suitable t , we have

$$\nabla r \approx \delta_t(\nabla s) \quad (29)$$

then the problem (26) can be defined as

$$\hat{r} = \arg \min_r \int_{\Omega} \sqrt{(\nabla_x r - \tau_t(\nabla_x s))^2 + (\nabla_y r - \tau_t(\nabla_y s))^2} \quad (30)$$

The L¹ Retinex model is also implemented using split Bregman method by introducing $\mathbf{d} = (d_x, d_y)$ as

$$\begin{aligned} d_x &= \nabla_x r - \tau_t(\nabla_x s) \\ d_y &= \nabla_y r - \tau_t(\nabla_y s) \end{aligned} \quad (31)$$

then equation (26) can be rewritten as

$$(\hat{r}, \hat{\mathbf{d}}) = \arg \min_{r,d} \int_{\Omega} \sqrt{d_x^2 + d_y^2} \text{ s.t. } \mathbf{d} = \nabla r - \delta_t(\nabla s) \quad (32)$$

then the Bregman iteration can be written as

$$\begin{aligned} r^{k+1} &= \arg \min_r -\lambda \langle \text{div} \mathbf{b}^k, r \rangle + \frac{\lambda}{2} \|\mathbf{d}^k - \nabla r + \delta_t(\nabla s)\|_2^2 \\ \mathbf{d}^{k+1} &= \arg \min_d \int_{\Omega} |\mathbf{d}| - \lambda \langle \mathbf{b}^k, \mathbf{d} \rangle + \frac{\lambda}{2} \|\mathbf{d} - \nabla r^{k+1} + \delta_t(\nabla s)\|_2^2 \\ \mathbf{b}^{k+1} &= \mathbf{b}^k - \mathbf{d}^{k+1} + \nabla r^{k+1} - \delta_t(\nabla s) \end{aligned} \quad (33)$$

The computation procedure of subproblems (33) is as well as TV's, so that won't be covered again here.

E. High-order Total Variational L^1 model (HoTVL¹)

As we know, the total variation method [44] may cause the so called staircase artifact phenomenon. In order to eliminate this effect, a number of higher order functions have been proposed such as infimal convolution of total variation (ICTV) [45]-[47] and total generalized variation (TGV) [48], which are defined as

$$\text{ICTV}_\beta(u) = \inf_{u=v+w} \int_\Omega |\nabla v| dx + \beta \int_\Omega |\nabla^2 w| dx \quad (34)$$

where $\int_\Omega |\nabla^2 w|$ is the total variation of Hessian of w for $w \in W^{2,1}(\Omega)$.

$$\text{TGV}_\beta^2(u) = \inf_w \int_\Omega |\nabla u - w| dx + \beta \int_\Omega |\nabla w| dx \quad (35)$$

where $\nabla w \in \mathbb{R}^{2 \times 2}$ denotes the gradient of the deformation field $w \in \mathbb{R}^2$.

Liang and Zhang [49] proposed the first higher order TV + L^1 variational model decomposition as

$$\min_{r,l} \left\{ \begin{aligned} \varepsilon_{\alpha,\beta}(r,l) &= \frac{1}{2} \int_\Omega (s-r-l)^2 dx \\ &+ \alpha \left(\int_\Omega |\nabla r| \right) dx + \beta \int_\Omega |\nabla^2 l| dx \end{aligned} \right\} \quad (36)$$

Let $-u = r + l$, then the above formula is equivalent to

$$\begin{aligned} \min_{u,l} \left\{ \frac{1}{2} \int_\Omega (u-s)^2 dx + \alpha \int_\Omega (|\nabla(u-l)| + \beta |\nabla^2 l|) dx \right\} \\ \Leftrightarrow \min_u \left\{ \frac{1}{2} \int_\Omega (u-s)^2 dx + \alpha \text{ICTV}_\beta(u) \right\} \end{aligned} \quad (37)$$

Further replace ∇l by $\mathbf{v} = [v_1, v_2]^T$ in equation (36), then the following model can be obtained

$$\begin{aligned} \min_{u,v} \left\{ \frac{1}{2} \int_\Omega (u-s)^2 dx + \alpha \int_\Omega (|\nabla u - \mathbf{v}| + \beta |\nabla \mathbf{v}|) dx \right\}, \\ \Leftrightarrow \min_u \left\{ \frac{1}{2} \int_\Omega (u-s)^2 dx + \alpha \text{TGV}_\beta^2(u) \right\} \end{aligned} \quad (38)$$

To overcome some drawbacks of model (36) and infimal convolution (37), (38), an extended version of model (36) is presented as

$$\min_{r \in \mathcal{B}_r, l \in \mathcal{B}_l} \left\{ \begin{aligned} \varepsilon_{\alpha,\beta,\tau}(r,l) &= \frac{1}{2} \int_\Omega (s-r-l)^2 dx \\ &+ \alpha \left(\int_\Omega |\nabla r| \right) dx + \beta \int_\Omega |\nabla^2 l| dx + \frac{\tau}{2} \int_\Omega l^2 dx \end{aligned} \right\} \quad (39)$$

where τ is a small positive parameter to ensure the boundedness of l , \mathcal{B}_r and \mathcal{B}_l are the box constraint for reflectance r and illumination l respectively.

The problem of (39) can be solved by primal dual splitting [50], [51] and split Bregman algorithms [43] with operator splitting techniques. In this paper, the split inexact Uzawa (SIU) method presented in [52] which is an inexact variant of alternating direction method of multipliers (ADMM) [53] adopted for this problem.

After define the following formulas

$$H(x) = \frac{1}{2} \|s - Ax\|^2 \quad (40)$$

$$J(x, y) = \alpha \|y\|_{1,\beta} + t\mathcal{B}(x) + \frac{\tau}{2} \|Bx\|^2$$

where

$$x = \begin{bmatrix} r \\ l \end{bmatrix}, y = \begin{bmatrix} u \\ v \end{bmatrix}, u = \nabla r, v = \nabla^2 l, A = [\text{Id}, \text{Id}], B = [0, \text{Id}],$$

$$\|y\|_{1,\beta} = \|u\|_1 + \beta \|v\|_1,$$

then (39) can be formulated into the following form

$$\min_{x,y} H(x) + J(x, y) \quad \text{s.t.} \quad Lx = y \quad (41)$$

whose augmented Lagrangian formula is

$$\begin{aligned} \max_p \min_{x,y} \{ \mathcal{L}(p; x, y) = H(x) + J(x, y) \\ + \langle p, Lx - y \rangle + \frac{\nu}{2} \|Lx - y\|^2 \} \end{aligned} \quad (42)$$

where $L = \begin{bmatrix} \nabla, 0 \\ 0, \nabla^2 \end{bmatrix}$ and $p = \begin{bmatrix} p_r \\ p_l \end{bmatrix}$ is the Lagrangian multiplier.

Then the SIU iteration can be written as

$$\begin{cases} x^{k+1} = \arg \min_x \mathcal{L}(p^k; x, y^k) + \frac{1}{2} \|x - x^k\|_{M_\nu} \\ y^{k+1} = \arg \min_y \mathcal{L}(p^k; x^{k+1}, y) \\ p^{k+1} = p^k + \nu(Lx^{k+1} - y^{k+1}) \end{cases} \quad (43)$$

where $M_\nu = \text{Id} - \nu L^T L$, which is a positive definite matrix.

The update of x^{k+1} and y^{k+1} is presented respectively as follows.

$$\begin{aligned} x^{k+1} &= \arg \min_x t\mathcal{B}(x) + \frac{1}{2} \|s - Ax\|^2 \\ &+ \frac{\tau}{2} \|Bx\|^2 + \frac{1}{2} \|x - w^k\|^2 \end{aligned} \quad (44)$$

where $w^k = x^k - L^T(\nu Lx^k + p^k - \nu y^k)$.

$$\begin{aligned} y^{k+1} &= \arg \min_y \alpha \|y\|_{1,\beta} + \frac{\nu}{2} \|y - Lx^{k+1} - p^k / \nu\|^2 \\ &= \arg \min_y \alpha \|u\|_1 + \frac{\nu}{2} \|u - \nabla r^{k+1} - p_r^k / \nu\|^2 \\ &+ \alpha \beta \|v\|_1 + \frac{\nu}{2} \|v - \nabla^2 l^{k+1} - p_l^k / \nu\|^2 \end{aligned} \quad (45)$$

Finally, the update for the dual variable p^{k+1} is straightforward.

IV. NONLOCAL VARIATIONAL RETINEX MODELS

In this section, we recall and give a few definitions of nonlocal (NL) differential operators proposed by Gilboa and Osher [54] that can be viewed as an extension of spectral graph theory and the diffusion geometry framework to functional analysis and PDE-like evolutions.

A. Basic Definitions (Differential Operators)

First, the definitions of particular products and norms of scalars and nonlocal vectors are given as follows.

Definition 4.1. To begin, we extend the notion of NL partial derivatives by the following definition:

$$\partial_y u(x) := \frac{u(y) - u(x)}{\widehat{d}(x, y)} \quad (46)$$

where $x, y \in \Omega$, $0 < \widehat{d}(x, y) \leq \infty$, and \widehat{d} is a positive measure defined between points x and y .

Definition 4.2. To keep with standard notations related to regularization framework on graphs which uses similar operators [55], we define the weights as

$$w(x, y) = \hat{d}^{-2}(x, y), \quad 0 \leq w(x, y) < \infty \quad (47)$$

Definition 4.3. In general, the weights are assumed symmetric, that is $w(x, y) = w(y, x)$. The NL derivation (46) can be written as

$$\partial_y u(x) := (u(y) - u(x)) \sqrt{w(x, y)} \quad (48)$$

Definition 4.4. The NL gradient $\nabla_w u(x) : \Omega \rightarrow \Omega \times \Omega$ is defined as the vector of all partial derivatives:

$$(\nabla_w u)(x, y) := (u(y) - u(x)) \sqrt{w(x, y)} \quad (49)$$

Definition 4.5. Vectors \vec{v} is denoted as $\vec{v} = v(x, y) \in \Omega \times \Omega$. The standard L^2 inner product is used for functions

$$\langle u_1, u_2 \rangle := \int_{\Omega} u_1(x) u_2(x) dx \quad (50)$$

Definition 4.6. Then the dot product for vectors at x is defined

$$\langle \vec{v}_1, \vec{v}_2 \rangle(x) := \int_{\Omega} v_1(x, y) v_2(x, y) dy \quad (51)$$

and an inner product is

$$\langle \vec{v}_1, \vec{v}_2 \rangle := \langle \vec{v}_1 \cdot \vec{v}_2, 1 \rangle = \int_{\Omega \times \Omega} v_1(x, y) v_2(x, y) dx dy \quad (52)$$

Definition 4.7. The Norm of a vector is

$$|\vec{v}|(x) := \sqrt{\vec{v}_1 \cdot \vec{v}_2} = \sqrt{\int_{\Omega} v(x, y)^2 dy} \quad (53)$$

With the above inner products the NL divergence $\text{div}_w \vec{v}(x) : \Omega \times \Omega \rightarrow \Omega$ is defined as the adjoint of the NL gradient:

$$\langle \nabla_w u, \vec{v} \rangle = \langle u, -\text{div}_w \vec{v} \rangle \quad (54)$$

Definition 4.8. The expression for the divergence is easily found as

$$\text{div}_w \vec{v}(x) := \int_{\Omega} (v(x, y) - v(y, x)) \sqrt{w(x, y)} dy \quad (55)$$

Definition 4.9. The nonlocal Laplacian can now be defined by

$$\begin{aligned} \Delta_w u(x) &:= \frac{1}{2} \text{div}_w (\nabla_w u(x)) \\ &= \int_{\Omega} (u(y) - u(x)) w(x, y) dy \end{aligned} \quad (56)$$

Note that in order to get the standard Laplacian definition which relates to the graph Laplacian we need a factor of 1/2.

Definition 4.10. The NL (mean) curvature can be formulated:

$$\begin{aligned} k_w &:= \text{div}_w \left(\frac{\nabla_w u}{|\nabla_w u|} \right) \\ &= \int_{\Omega} (u(y) - u(x)) w(x, y) \left(\frac{1}{|\nabla_w u|(x)} + \frac{1}{|\nabla_w u|(y)} \right) dy \end{aligned} \quad (57)$$

with norm of NL gradient at q :

$$|\nabla_w u|(q) := \sqrt{\int_{\Omega} (u(z) - u(q))^2 w(q, z) dz} \quad (58)$$

B. Nonlocal Retinex

Zosso and Osher [56] presented a unifying framework of nonlocal Retinex models. They summarize an L^p NL Retinex functionals:

$$\hat{r} = \arg \min_r \left\{ \|\nabla_w r - \nabla_w r^s\|_p^p + \alpha \|r\|_2^2 + \beta \|r - s\|_2^2 \right\} \quad (59)$$

Here, we first focus on functionals with a TV-type regularizer and their numerical optimization, then we will introduce several representative NL Retinex models.

NL TV Retinex

Following Buades, Coll, and Morel's nonlocal mean method [57], Gilboa and Osher introduced the nonlocal TV regularizer [54]. This is another successful method in image processing, especially for textured images [58].

For an image $u : \Omega \rightarrow \mathbb{R}$, we can define the nonlocal weight between two pixel x and y :

$$w_h(x, y) = \exp \left\{ \frac{-G_a * (u(x) - u(y))^2}{2h^2} \right\} \quad (60)$$

where G_a is the Gaussian kernel with standard deviation a . With the nonlocal gradient operator as shown in definition 4.4, the nonlocal TV regularizer can be defined as

$$\int_{\Omega} |\nabla_w u| = \int_{\Omega} \left(\int_{\Omega} (u(y) - u(x))^2 w(x, y) dy \right)^{\frac{1}{2}} dx \quad (61)$$

So the nonlocal TV regularized model for Retinex theory is

$$\hat{r} = \arg \min_r \left\{ \int_{\Omega} \left(t |\nabla_w r| + \frac{1}{2} \|\nabla(r - s)\|_2^2 \right) \right\} \quad (62)$$

Similarly, we have a numerical algorithm using the split Bregman method to the nonlocal TV regularized model by introducing an auxiliary variable $\mathbf{d} = \nabla_w r = \nabla_w u$. In discrete version, d can be written as $\mathbf{d} = (d_1, d_1, \dots, d_m)$ with $d_j = d_{j1}, d_{j2}, \dots, d_{jm} \in \mathbb{R}^n$. And the minimization problem is

$$\hat{r} = \arg \min_r t \sum_{k=1}^n \left(\sum_{j=1}^m d_{jk}^2 \right)^{\frac{1}{2}} + \frac{1}{2} \|D(r - s)\|_2^2 \quad \text{s.t. } \mathbf{d} = D_w r \quad (63)$$

When we plug this into the general form of the split Bregman iterations, we have

$$\begin{aligned} \mathbf{d}^{k+1} &= \arg \min_{\mathbf{d}} \frac{t}{\lambda} \sum_{k=1}^n \left(\sum_{j=1}^m d_{jk}^2 \right)^{\frac{1}{2}} \\ &\quad - \langle \mathbf{b}^k, \mathbf{d} - \mathbf{d}^k \rangle + \frac{1}{2} \|\mathbf{d} - D_w r^k\|_2^2 \\ r^{k+1} &= \arg \min_u \frac{1}{2\lambda} \|D(r - s)\|_2^2 \\ &\quad + \langle \mathbf{b}^k, D_w(r - r^k) \rangle + \frac{1}{2} \|\mathbf{d}^{k+1} - D_w r\|_2^2 \\ \mathbf{b}^{k+1} &= \mathbf{b}^k - (\mathbf{d}^{k+1} - D_w r^{k+1}) \end{aligned} \quad (64)$$

We explicitly solve the problem (64) and get the nonlocal TV Bregman iterative algorithm.

Initialize $u^0 = s, \mathbf{d}^0 = \mathbf{b}^0 = 0$

While $\|u^{k+1} - u^k\|_2 / \|u^{k+1}\|_2 > \varepsilon$

$$(1) \mathbf{d}^{k+1} = \text{shrink}_{t/\lambda} (D_w r^k + \mathbf{b}^k) \quad (65)$$

$$(2) r^{k+1} = (D^T D + \lambda D_w^T D_w)^{-1} (\lambda D_w^T (\mathbf{d}^{k+1} - \mathbf{b}^k) + D^T D s)$$

$$(3) \mathbf{b}^{k+1} = \mathbf{b}^k - (\mathbf{d}^{k+1} - D_w r^{k+1})$$

end While

Here shrink_t is the nonlocal isotropic shrinkage function. For $v \in \mathbb{R}^m$, denote $v = (x_1, x_2, \dots, x_m)^T$ and $x_i = (v_{in+1}, v_{in+2}, \dots, v_{(i+1)n})$.

$$\text{shrink}_t(v_{in+1}) = \begin{cases} 0 & \text{if } \|x_i\|_2 \leq t \\ v_{in+1} - \frac{t}{\|x_i\|_2} v_{in+1} & \text{if } \|x_i\|_2 > t \end{cases} \quad (66)$$

L^2 -gradient-fidelity NL Retinex

The energy of the L^2 -gradient-fidelity nonlocal Retinex is

$$J(r) = \|\nabla_w r - \nabla_{w,f} s\|_2^2 + \alpha \|r\|_2^2 + \beta \|r - s\|_2^2 \quad (67)$$

The corresponding Euler-Lagrange equations are

$$2(-\Delta_w \hat{r} + \Delta_{w,f} i + \alpha \hat{r} + \beta \|\hat{r} - s\|) = 0 \quad (68)$$

And the estimate of reflectance \hat{r} can be recovered as

$$\hat{r} = ((\alpha + \beta)I - L)^{-1} (\beta s - \Delta_{w,f} s) \quad (69)$$

where I is the identity matrix and L is the Laplacian matrix derived from the weights $w(x, y)$:

$$L_{xy} = \begin{cases} w(x, y) + w(y, x), & x \neq y \\ -\sum_{z \neq x} w(x, z) + w(z, x), & x = y \end{cases} \quad (70)$$

As the graph Laplacian L is negative semidefinite, the operator $(\alpha + \beta)I - L$ is diagonally dominant. The problem \hat{r} can be solved either by a Gauss-Seidel algorithm or successive over relaxation (SOR) method.

L^1 -gradient-fidelity NL Retinex

The L^1 -based problem is written as

$$\min_r \left\{ \|\nabla_w r - \nabla_{w,f} s\|_1 + \alpha \|r\|_2^2 + \beta \|r - s\|_2^2 \right\} \quad (71)$$

which can be split into

$$\min_{r,e} \left\{ \|e - \nabla_{w,f} s\|_1 + \alpha \|r\|_2^2 + \beta \|r - s\|_2^2 \right\} \text{ s.t. } e = \nabla_w r \quad (72)$$

The constraint can be addressed by applying the following augmented Lagrangian, including a quadratic penalty and a Lagrangian multiplier term:

$$\begin{aligned} AL(r, e, \mu) = & \|e - \nabla_{w,f} s\|_1 + \alpha \|r\|_2^2 \\ & + \beta \|r - s\|_2^2 + \rho \|\nabla_w r - e\|_2^2 + 2\langle \mu, \nabla_w r - e \rangle \end{aligned} \quad (73)$$

After that, the above L^1 -minimization problem can be addressed by applying the ADMM algorithm, which includes three steps:

Step 1. Solving the L^2 -minimization in r .

$$\begin{aligned} r^{k+1} = & \arg \min_r \left\{ \alpha \|r\|_2^2 + \beta \|r - s\|_2^2 + \rho \|\nabla_w r - e^k + \mu^k / \rho\|_2^2 \right\} \\ = & ((\alpha + \beta)I - \rho L)^{-1} (\beta s - \rho \text{div}_w (e^k - \mu^k / \rho)) \end{aligned} \quad (74)$$

Step 2. Shrinkage of e .

$$\begin{aligned} e^{k+1} = & \arg \min_r \left\{ \|e - \nabla_{w,f} s\|_1 + \rho \|e - \nabla_w r^{k+1} - \mu^k / \rho\|_2^2 \right\} \\ = & S_{1/2\rho}^s (\nabla_w r^{k+1} - \nabla_{w,f} s + \mu^k / \rho) + \nabla_{w,f} s \end{aligned} \quad (75)$$

Step 3. Updating the Lagrangian multiplier μ .

$$\mu^{k+1} = \mu^k + \rho (\nabla_w r^{k+1} - e^{k+1}) \quad (76)$$

The update of r^{k+1} in step 1 is the most time consuming part, this process can be speeded up using the split Bregman method or the Gauss-Seidel or conjugate gradient method.

L^0 -gradient-fidelity NL Retinex

Finally, the nonconvex L^0 -based gradient fidelity optimization problem is written as

$$\min_r \left\{ \|\nabla_w r - \nabla_{w,f} s\|_0 + \alpha \|r\|_2^2 + \beta \|r - s\|_2^2 \right\} \quad (77)$$

To solve this difficulty problem, here, it is treated similarly to L^1 minimization problem, which also first split into a linearly constrained minimization functional as follows.

$$\min_{r,e} \left\{ \|e - \nabla_{w,f} s\|_0 + \alpha \|r\|_2^2 + \beta \|r - s\|_2^2 \right\} \text{ s.t. } e = \nabla_w r \quad (78)$$

Now the suboptimization problems are easy to solve, but it is not overall convergence. This problem can be addressed in analogy to the L^1 -gradient fidelity in subsection of “ L^1 -gradient-fidelity NL Retinex”. Here, both a quadratic penalty and a Lagrangian multiplier are used to solve iteratively along each direction. The mainly difference is one more step is added to the iterative process, which is used to enforce convergence. The numerical implementation process is as follow.

$$\begin{aligned} r^{k+1} = & \arg \min_r \left\{ \alpha \|r\|_2^2 + \beta \|r - s\|_2^2 + \rho^k \|\nabla_w r - e^k + \mu^k / \rho^k\|_2^2 \right\} \\ = & ((\alpha + \beta)I - \rho^k L)^{-1} (\beta s - \rho^k \text{div}_w (e^k - \mu^k / \rho^k)) \\ e^{k+1} = & \arg \min_r \left\{ \|e - \nabla_{w,f} s\|_0 + \rho^k \|e - \nabla_w r^{k+1} - \mu^k / \rho^k\|_2^2 \right\} \\ = & S_{1/\sqrt{\rho^k}}^h (\nabla_w r^{k+1} - \nabla_{w,f} s + \mu^k / \rho^k) + \nabla_{w,f} s \\ \mu^{k+1} = & \mu^k + \rho^k (\nabla_w r^{k+1} - e^{k+1}) \\ \rho^{k+1} = & \rho^k \cdot s \end{aligned} \quad (79)$$

V. NUMERICAL RESULTS

In this section, the performance of different variational Retinex methods is presented and compared quantitatively and qualitatively. To easily discriminate the difference of these variational Retinex models, their energy functionals are summarised in Table I. All experiments are executed using Matlab 2012b on a Windows 7 platform with an Intel H81 CPU i5-4460 at 3.20GHz and 4GB memory.

TABLE I VARIATIONAL RETINEX MODELS

Model	Energy functional
Kimmel's	$E(l) = \int_{\Omega} \nabla l ^2 + \alpha \int_{\Omega} (l - s)^2 + \beta \int_{\Omega} \nabla(l - s) ^2$
TV	$E(r) = \int_{\Omega} t \nabla r + \frac{1}{2} \int_{\Omega} \nabla r - \nabla s ^2$
TV-L ²	$E(r, l) = \int_{\Omega} Dr + \frac{\alpha}{2} \int_{\Omega} \nabla l ^2 + \frac{\beta}{2} \int_{\Omega} (l - r - s)^2 + \frac{\mu}{2} \int_{\Omega} l^2$
L ¹	$E(r) = \int_{\Omega} \nabla r - \delta_i(\nabla s) $
HoTVL ¹	$E(r, l) = \frac{1}{2} \int_{\Omega} (s - r - l)^2 + \alpha (\int_{\Omega} \nabla r) + \beta \int_{\Omega} \nabla^2 l + \frac{\tau}{2} \int_{\Omega} l^2$
NL-TV	$E(r) = t \int_{\Omega} \nabla_w r + \frac{1}{2} \int_{\Omega} \ \nabla(r - s)\ _2^2$
NL-L ^p (P=0,1,2)	$E(r) = \int_{\Omega} \ \nabla_w r - \nabla_{w,f} s\ _p + \alpha \int_{\Omega} \ r\ _2^2 + \beta \int_{\Omega} \ r - s\ _2^2$

A. Qualitative Measures

In our experiments, we apply the numerical algorithms to several classic test images. Fig. 3 lists all the color images and grayscale images which will be used in the following experiments. These testing images suffer of one or more of the following problems such as low contrast, non-uniform lighting, blurring, color diminished and noise. The experimental comparison results demonstrate the effectiveness of the proposed models with their algorithms.

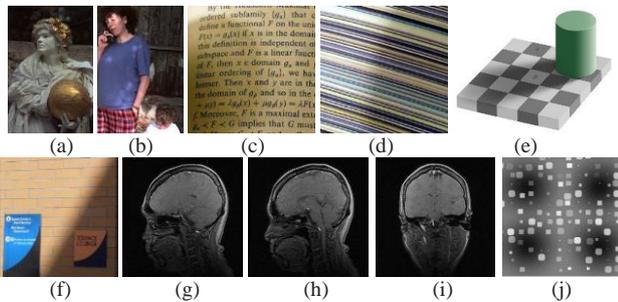


Fig. 3. The original tested images: (a)-(j) used for the following experiments.

We apply the Kimmel’s process, with optimal parameters $\alpha = 0.0001$ $\beta = 0.1$ and fixed $\gamma = 3$. Fig. 4 demonstrates the influence of the α and β values on the reconstructed illumination and reflectance images. The α values change from $1e-7$ to $1e-1$ and β values change from $1e-4$ to 0 with relatively minor effect on the output quality. The results present a good image quality and parameter robustness. It was shown that for a wide range of the involved parameters, the output quality is practically the same.

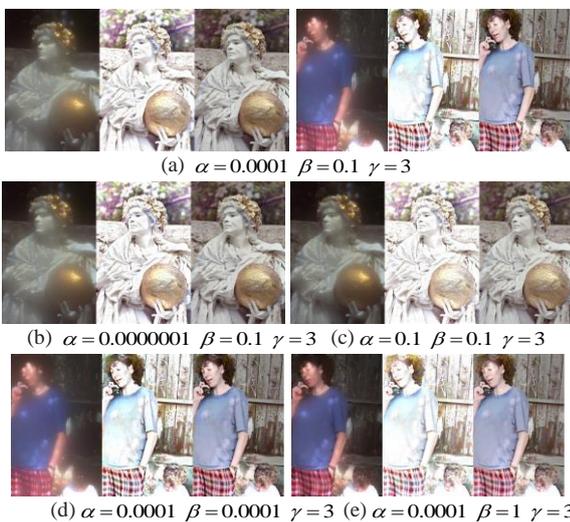


Fig. 4. Performance of Kimmel’s model, illumination image, reflectance image, output image (left to right of each synthetic images); results with optimal parameters $\alpha = 0.0001$ $\beta = 0.1$ in (a); influence of α with constant β in (b) and (c); influence of β with constant α in (d) and (e).

In TV model’s experiment [40], one testing image is a piece of text in a shadow of some object shown in Fig. 3(c), another is a piece of cloth with colorful bands and each band is of a constant color, as shown in Fig. 3(d). To better demonstrate the performance of TV regularized model, we compare it with PDE-based model and NL TV regularized model. From the resulting image of Fig. 5, it can be seen that

both the NL TV regularized model and TV regularized model outperform PDE-based model. The shadow in the resulting image of NL TV regularized model is almost disappeared and each color band is constant as shown in Fig. 5(c) and 5(f), especially in the treatment of the texture images. Meanwhile, comparing with the PDE-based model, in the resulting image of Fig. 5(d) and 5(e) of TV model, the shadow is less obvious and the color of the background is more constant. It needs to be noted that if the parameter t is selected larger, the effect of the shadow is weaker.

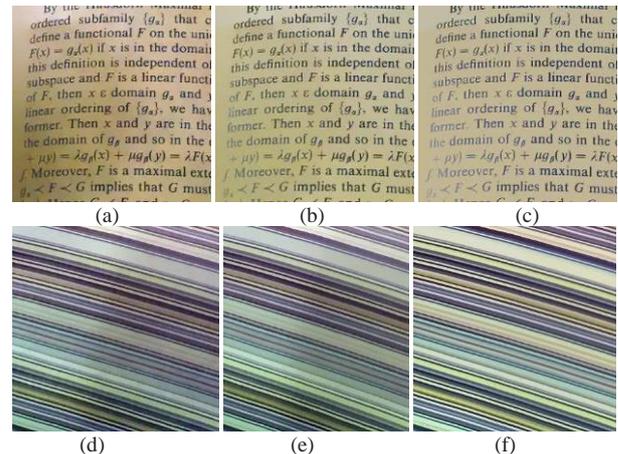


Fig. 5. Performance of TV model: The resulting images of the PDE-based algorithm with $t=20$ in (a) and (d); The resulting images of TV regularized model with $t=15$ in (b) and (e); The result of images of NL TV regularized model with $t=15$ in (c) and (f).

Fig. 6 presents the testing examples of L^1 Retinex model [42], one is the image “Adelson’s checker shadow illusion” shown in Fig. 6(a) and (b), it can be seen that that the contrast in the recovered reflectance of the L^1 Retinex model is stronger than the PDE model. It is also shown that a larger number of reflectance information is preserved in the recovered illumination image of the PDE model. On the contrary, the illumination image from the L^1 Retinex method takes the least information of the reflectance and gives a better result. Fig. 6(c)-(e) are the recovered results of ‘wall in shadow’ from L^1 Retinex model with different values of t , it demonstrates that the illumination (shadow effect) is eliminated gradually as t increases.

The L^1 Retinex model also makes a good performance on medical images such as MRI in Fig. 3(g)-(j). Due to the apparent bias field effect in the original clinical images, it can be hardly observed the area of the neck. Fig. 6(f)-(h) are the recovered MRI from the L^1 Retinex model, it is shown that the details near the lower neck area are significantly visible comparing with the original images.

To demonstrate the performance of HoTVL¹ model [49], two image decomposition problems are considered, including synthetic example, Retinex illusion example. We also further compare HoTVL¹ model with two other variational models (TV) and (L^1 Retinex) proposed in [24] and [42] respectively.

We start with synthetic example. As we can observe, L^1 Retinex and HoTVL¹ produce better visual results than TV model. From TV model’s result, more information of the edges of reflectance is contained in the recovered illumination l . Conversely, only a little information of reflectance r is contained in the recovered illumination l after applying L^1

Retinex and HoTVL¹ models. Similar to the synthetic example, in Retinex illusion, the L¹ Retinex and HoTVL¹ models provide visually preferable results compare to TV model, however, the recovered illumination in HoTVL¹ model contains less reflectance information than L¹ Retinex.

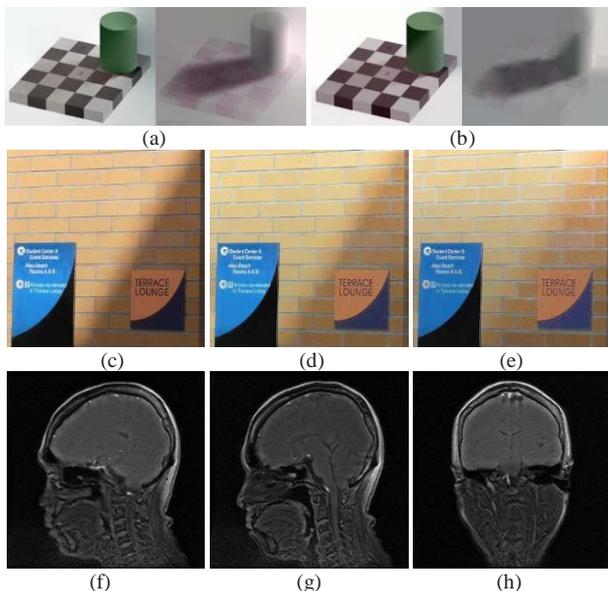


Fig. 6. Performance of L¹ Retinex model: recovered reflectance and illumination from the PDE method in (a); recovered reflectance and illumination from the L¹ Retinex model in (b); influence of threshold t : (c) $t=5$, (d) $t=10$, (e) $t=15$; corrected clinical images by the L¹ Retinex model in (f), (g), (h).

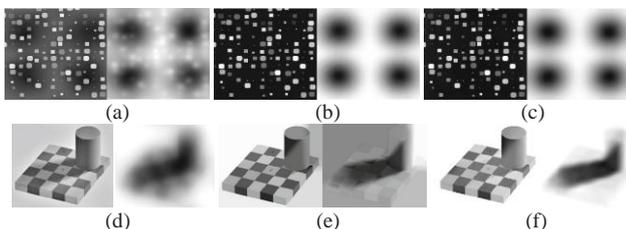


Fig. 7. Decomposition comparison of two images: recovered reflectance r and illumination l of synthetic example (a-c) and Retinex illusion example (d-f) by TV, L¹ Retinex and HoTVL¹ models, respectively.

B. Quantitative Measures

In this section, this study utilises five quantitative evaluation metrics namely root mean square error (RMSE), peak signal to noise ratio (PSNR), structural similarity (SSIM) [59], discrete entropy and contrast measure (DECM) defined by Celik [60], and the CPU time to evaluate the performance of these methods. To better illustrate the performance of these variational Retinex models, five representative images from Fig. 3(a) to Fig. 3(e) and 300 test images from UC. Berkeley image data set are selected to execute on.

The RMSE and PSNR are normally used to assess the noise, lower RMSE and higher PSNR values indicate less noise. Their average values in Fig. 8 and Table II shows that HoTVL¹ and NL-TV models perform better than most of the other existing models. For some images, the HoTVL¹ obtains Lower RMSE values and higher PSNR values than NL-TV model. However, NL-TV model achieves Lower average RMSE and higher average PSNR. The SSIM and DECM values (values are both between 0 (worst) and 1 (best)) shown

in Fig. 9 and Table III also indicate that HoTVL¹ and NL-TV models improve the global discrete entropy and local contrast measures better than others'. It is shown that the SSIM values obtained from these models are both more than 0.85 and the average value of NL-TV is 0.9128, which produce satisfactory results.

To conclude this part, we further present the computational CPU time comparison of the above models. In addition to the aforementioned split Bregman algorithm and the alternating direction method of multipliers (ADMM) algorithm, there also exist some other approaches, such as the augmented Lagrangian algorithm [61], [62], the primal-dual algorithm [63], [64] and etc., to solve the problems of minimising the energy functionals. In the experiment, we apply the representative split Bregman for each model and compare their computational speed. Tables IV demonstrates computational efficiency of different models. For all models, we set stopping criteria as $|E^k - 1|/E^k \leq \varepsilon$, where E is the value of energy functional of each model, and ε is a small tolerance used to stop iteration. Here, $\varepsilon = 10^{-5}$ in all cases. Table IV illustrates NL TV is the slowest model around 10s, while TV is the fastest model. Furthermore, L¹ and HoTVL¹ models are less efficient than Kimmel's and TV models. However, the total iterations of NL TV model are fewest among these models, and L¹ and HoTVL¹ models are much more than others'.

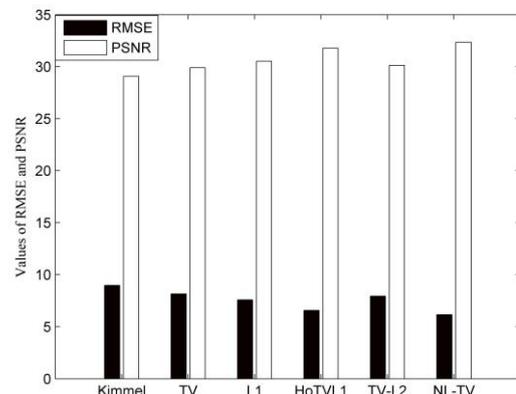


Fig. 8. Comparison of variational Retinex models in average values of RMSE and PSNR of Fig. 3(a)-Fig. 3(e).

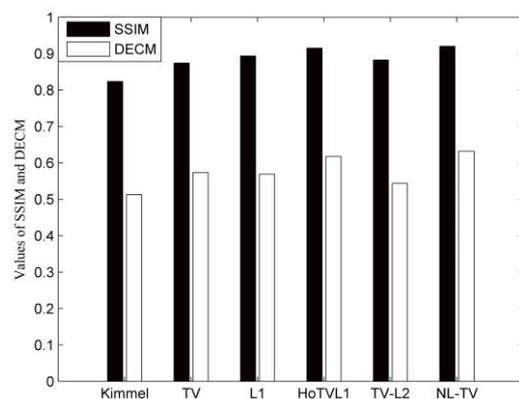


Fig. 9. Comparison of variational Retinex models in average values of SSIM and DECM of Fig. 3(a)-Fig. 3(e).

TABLE II
QUANTITATIVE COMPARISON IN RMSE AND PSNR (THE BOLD VALUES EXPRESS THE BEST METRIC VALUES)

Model	RMSE (Average)	PSNR (Average)
Kimmel's	9.8905	28.2264
TV	9.1348	28.9168
L ¹	7.9275	30.1481
HoTVL ¹	6.6251	31.7069
TV-L ²	8.5346	29.5071
NL-TV	6.0781	32.4554

TABLE III
QUANTITATIVE COMPARISON IN SSIM AND DECM (THE BOLD VALUES EXPRESS THE BEST METRIC VALUES)

Model	SSIM (Average)	DECM (Average)
Kimmel's	0.8632	0.5078
TV	0.8803	0.5641
L ¹	0.8901	0.5793
HoTVL ¹	0.9016	0.6072
TV-L ²	0.8924	0.5316
NL-TV	0.9128	0.6425

TABLE IV
QUANTITATIVE COMPARISON IN CPU TIME OF PER ITERATION AND TOTAL ITERATIONS (THE BOLD VALUES EXPRESS THE BEST METRIC VALUES)

Model	Time (s)	Total iterations
Kimmel's	0.0523	NA
TV	0.0357	80
L ¹	0.1972	175
HoTVL ¹	1.2406	100
TV-L ²	0.1354	160
NL-TV	10.1423	35

VI. CONCLUSION

In this paper, several variational Retinex models are introduced, and their application for image enhancement is illustrated. We present the detailed discretisation process based on discrete finite different scheme and numerical implementation of split Bregman algorithm, alternating direction method of multipliers (ADMM) algorithm and fast Fourier transform (FFT) to solve the problems. Further, the advantages and disadvantages of these models are demonstrated by extensive comparative experiments in last section.

Among all the models, in Kimmel's model, the reflection function is not considered. The TV regularized model is

efficient to recover piecewise constant images due to property of TV regularizer, but it usually loses information about reflectance. L¹ variational Retinex model is based on minimizing an L¹ norm, which ensures that both the recovered reflectance and illumination have better quality than previous works. Besides, this model can be also applied to shadow elimination problem and MRI and hyperspectral images. HoTVL¹ model extracts preferable illumination, and detail preserved reflectance, comparing to TV and L¹ Retinex decomposition models. Nonlocal variational models can achieve desirable enhancement result, especially texture image. However, the Nonlocal and HoTVL¹ models are more time-consuming as result of computational complexity of algorithms.

REFERENCES

- [1] P. Felzenszwalb, D. Huttenlocher, "Efficient graph-based image segmentation," *Int'l J. Computer Vision*, vol. 59, no. 2, pp. 167-181, 2004.
- [2] N. M. Saad, S.A.R. A.-Bakar, S. Muda, M. Mokji, and A.R. Abdullah, "Fully automated region growing segmentation of brain lesion in diffusion-weighted MRI," *IAENG International Journal of Computer Science*, vol. 39, no. 2, pp. 155-164, 2012.
- [3] M. F. Hashmi, A. R. Hambarde, A. G. Keskar, "Robust image authentication based on HMM and SVM classifiers," *Engineering Letters*, vol. 22, no. 4, pp. 183-193, 2014.
- [4] K. He, X. Zhang, S. Ren, J. Sun, "Deep residual learning for image recognition," In: *IEEE Conference on Computer Vision and Pattern Recognition 2016*, pp. 770-778.
- [5] Z. Karel, "Contrast limited adaptive histogram equalization," in *Graphic Gems IV*. San Diego, CA, USA: Academic, pp. 474-485, 1994.
- [6] Y. T. Kim, "Contrast enhancement using brightness preserving bi-histogram equalization," *IEEE Trans. Consumer Electron.*, vol. 43, no. 1, pp. 1-8, 1997.
- [7] K. Iqbal, R. Abdul Salam, A. Osman, and A. Zawawi Talib, "Underwater image enhancement using an integrated color model," *IAENG International Journal of Computer Science*, vol. 34, no. 2, pp. 239-244, 2007.
- [8] Z. Y. Chen, B. R. Abidi, D. L. Page, and M. A. Abidi, "Gray-level grouping (GLG): An automatic method for optimized image contrast enhancement-part I: The basic method," *IEEE Trans. Image Process.*, Vol. 15, No. 8, pp. 2290-2302, 2006.
- [9] L. Tao and V.K. Asari, "Adaptive and integrated neighborhood-dependent approach for nonlinear enhancement of color images," *Journal of Electronic Imaging*, vol. 14, no. 4, pp. 1-14, 2005.
- [10] E. H. Land and J.J. McCANN, "Lightness and retinex theory," *J. Opt. Soc. Am.*, vol. 61, no. 1, pp. 1-11, 1971.
- [11] E. Provenzi, L. D. Carli, A. Rizzi, and D. Marini, "Mathematical definition and analysis of the retinex algorithm," *J. Opt. Soc. Amer. A.*, vol. 22, No. 12, pp. 2613-2621, 2005.
- [12] E. Provenzi, M. Fierro, A. Rizzi, L. De Carli, D. Gadia, and D. Marini, "Random spray retinex: A new retinex implementation to investigate the local properties of the model," *IEEE Trans. Image Process.*, vol. 16, no. 1, pp. 162-171, 2007.
- [13] J. A. Frankle, and J. J. McCann, "Method and apparatus for lightness imaging," *U.S. Patent 4384336 A*, 1983.
- [14] J. J. McCann, "Lessons learned from mondrians applied to real images and color gamuts," in *Proc. 7th Color Imag. Conf.*, pp. 1-8, 1999.
- [15] B. Funt, F. Ciurea, and J. McCann, "Retinex in matlabm," *J. Electron. Imaging*, vol. 13, pp. 48-57, 2004.
- [16] D. J. Jobson, Z. Rahman, and G. A. Woodell, "Properties and performance of the center/surround Retinex," *IEEE Trans. Image Process.*, vol. 6, no. 3, pp. 451-462, 1997.
- [17] Z. Rahman, D. J. Jobson, and G. A. Woodell, "Multi-scale retinex for color image enhancement," in *Proc. ICIP*, pp. 1003-1006, 1996.
- [18] Z. U. Rahman, D. J. Jobson and G. A. Woodell, "Retinex processing for automatic image enhancement," *J. Electron. Imag.*, vol. 13, No. 1, pp. 100-110, 2004.
- [19] J. M. Morel, A.-B. Petro, and C. Sbert, "A PDE formalization of retinex theory," *IEEE Trans. Image Process.*, vol. 19, pp. 2825-2837, 2010.

- [20] A. Blake, "Boundary conditions for lightness computation in mondrian world," *Comput. Vis. Graph. Imag. Process.*, vol. 32, pp. 314-327, 1985.
- [21] B. K. P. Horn, "Determining lightness from an image," *Comput. Graph. Imag. Process.*, vol. 3, no. 4, pp. 277-299, 1974.
- [22] R. Kimmel, M. Elad, D. Shaked, R. Keshet, and I. Sobel, "A variational framework for retinex," *Int. Journal of Computer Vision*, vol. 52, no. 1, pp. 7-23, 2003.
- [23] M. Elad, R. Kimmel, D. Shaked, and R. Keshet, "Reduced complexity retinex algorithm via the variational approach," *J. Vis. Commun. Imag. Represent.*, vol. 14, no. 4, pp. 369-388, 2003.
- [24] M. K. Ng and W. Wang, "A total variation model for Retinex," *SIAM J. Imaging Sci.*, vol. 4, pp. 345-365, 2011.
- [25] D. H. Choi, I. H. Jang, M. H. Kim, and N. C. Kim, "Color image enhancement based on single-scale retinex with a JND-based nonlinear filter," in: *ISCAS*, pp. 3948-3951, 2007.
- [26] D. H. Choi, I. H. Jang, M. H. Kim, and N. C. Kim, "Color image enhancement using single-scale retinex based on an improved image formation model," *16th European signal processing conference*, 2008.
- [27] G. L. Wen, S.G. Zheng, G. Liu, and S.K. Ning, "Image enhancement method based on the single-scale Retinex and color transferring," proceeding in: *International Conference on Electrical, Control and Automation (ICECA)*, pp. 645-651, 2014.
- [28] B. Jiang, G.A. Woodell and D.J. Jobson, "Novel multi-scale retinex with color restoration on graphics processing unit," *J. Real-Time Image Proc.*, pp. 1-15, 2014.
- [29] W. Wang, B. Li, J. Zheng, S. Xian, and J. Wang., "A fast multi-scale retinex algorithm for color image enhancement," *Int. Conf. on Wavelet Analysis and Pattern Recognition*, pp. 80-85, 2008.
- [30] S. Parthasarathy and P. Sankaran., "An automated multi scale retinex with color restoration for image enhancement," *2012 National Conference on Communications (NCC)*, Kharagpur, India, pp. 1-5, 2012.
- [31] D. H. Brainard and B. Wandell, "Analysis of the retinex theory of color vision," *J. Opt. Soc. Am. A*, vol. 3, pp. 1651-1661, 1986.
- [32] W. C. Zhang, X.J. An, S.D. Pan, "An improved recursive retinex for rendering high dynamic range photographs," in *Proceedings of the 2011 IEEE International Conference on Wavelet Analysis and Pattern Recognition*, pp. 121-125, 2011.
- [33] D. G. Hwang, W.R. Lee, Y.-J. Oh, and B.-M. Jun, "Frankle-McCann retinex by shuffling," in *Convergence and Hybrid Information Technology*, Springer-Verlag, Berlin, Heidelberg, pp. 381-388, 2012.
- [34] R. Sobol.: "Improving the retinex algorithm for rendering wide dynamic range photographs," *J. Electron. Imaging*, vol. 13, pp. 65-74, 2004.
- [35] E. Provenzi, L. De Carli, A. Rizzi, and D. Marini, "Mathematical definition and analysis of the retinex algorithm," *J Opt Soc Am A Opt Image Sci Vis*, vol. 22, no. 12, 2613-2621, 2005.
- [36] G. Gianini, A. Rizzi, and E. Damiani, "A retinex model based on absorbing markov chains," *Inf. Sci.*, pp. 149-174, 2016.
- [37] D. J. Jobson, Z. Rahman, and G. A. Woodell. "A multiscale retinex for bridging the gap between color images and the human observation of scenes," *IEEE Trans. on Image Processing*, vol. 6, no. 7, pp. 965-976, 1997.
- [38] J. M. Morel, A.B. Petro, and C. Sbert, "Fast implementation of color constancy algorithms," *Proc. SPIE*, 2009.
- [39] J. M. Morel, A.B. Petro, and C. Sbert, "A PDE formalization of retinex theory," *IEEE Trans. Image Process.*, vol. 19, no. 11, pp. 2825-2837, 2010.
- [40] W. Ma and S. Osher, "A TV bregman iterative model of retinex theory," UCLA CAM Report, UCLA, Los Angeles, CA, pp. 10-13, 2010.
- [41] M. K. Ng and W. Wang, "A total variation model for Retinex," *SIAM J. Imaging Sci.*, vol. 4, no. 1, pp. 345-365, 2011.
- [42] W. Ma, J.-M. Morel, S. Osher, and A. Chien, "An L^1 -based variational model for Retinex theory and its application to medical images," in *Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 153-160, 2011.
- [43] T. Goldstein and S. Osher, "The split bregman algorithm for $L1$ regularized problems," *SIAM J. Imaging Sci.*, vol. 2, no. 2, pp. 323-343, 2009.
- [44] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D*, vol. 60, no. 1, 259-268, 1992.
- [45] A. Chambolle, P.L. Lions, "Image recovery via total variation minimization and related problems," *Numer. Math.*, Vol. 76, No. 2, 167-188, 1997.
- [46] S. Setzer and G. Steidl, "Variational methods with higher order derivatives in image processing," In: N. Press (ed.) *Approximation XII*, Brentwood, pp. 360-386, 2008.
- [47] S. Setzer, G. Steidl, and T. Teuber, "Infimal convolution regularizations with discrete L^1 -type functionals," *Comm. Math. Sci.*, vol. 9, no. 3, pp. 797-872, 2011.
- [48] K. Bredies, K. Kunisch, and T. Pock, "Total generalized variation," *SIAM J. Imaging Sci.*, vol. 3, no. 3, pp. 492-526, 2010.
- [49] J. Liang, X. Zhang, "Retinex by higher order total variation L^1 decomposition," *Journal of Mathematical Imaging and Vision*, vol. 52, no. 3, pp. 345-355, 2015.
- [50] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *J. Math. Imaging Vis.*, vol. 40, no. 1, pp.120-145, 2011.
- [51] E. Esser, X. Zhang, and T.F. Chan, "A general framework for a class of first order primal-dual algorithms for convex optimization in imaging science," *SIAM J. Imaging Sci.*, pp. 1015-1046, 2010.
- [52] X. Zhang, M. Burger, and S. Osher, "A unified primal-dual algorithm framework based on Bregman iteration," *J. Sci. Comput.*, vol. 46, no. 1, pp. 20-46, 2011.
- [53] D. Gabay, "Applications of the method of multipliers to variational inequalities, Augmented Lagrangian methods: applications to the solution of boundary-value problems," 15, chap. IX, pp. 299-331, 1983.
- [54] G. Gilboa and S. Osher, "Nonlocal operators with applications to image processing," *Multiscale Model. Simul.*, vol. 7, pp. 1005-1028, 2008.
- [55] S. Bougleux, A. Elmoataz, and M. Melkemi, "Discrete regularization on weighted graphs for image and mesh filtering," in *Proceedings of the 1st International Conference on Scale Space and Variational Methods in Computer Vision (SSVM'07)*, *Lecture Notes in Comput. Sci.*, 4485, Springer-Verlag, Berlin, pp. 128-139, 2007.
- [56] D. Zosso, G. Tran, and S. J. Osher, "Non-local Retinex: a unifying framework and beyond," *SIAM J. Imag. Sci.*, vol. 8, no. 2, pp. 787-826, 2015.
- [57] A. Buades, B. Coll, and J.M. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Model. Simul.*, vol. 4, no. 2, 2005.
- [58] X. Zhang, M. Burger, X. Bresson, and S. Osher, "Bregmanized nonlocal regularization for deconvolution and sparse reconstruction," UCLA CAM Report 09-03, 2009.
- [59] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, "Image quality assessment: from error measurement to structural similarity," *IEEE T. Image Process.*, vol. 13, no. 1, pp. 600-612, 2004.
- [60] T. Celik, "Two-dimensional histogram equalization and contrast enhancement," *Pattern Recognit.*, vol. 45, no. 10, pp. 3810-3824, 2012.
- [61] X. C. Tai and C.L. Wu, "Augmented lagrangian method, dual methods and split bregman iteration for ROF model," *Scale Space and Variational Methods in Computer Vision Springer: Voss, Norway*, vol. 5567, pp. 502-513, 2009.
- [62] W. Zhu, X.C. Tai, and T.F. Chan, "Augmented lagrangian method for a mean curvature based image denoising model," *Inverse Problems and Imaging*, vol. 7, no. 4, pp. 1409-1432, 2013.
- [63] K. Bredies, K. Kunisch, and T. Pock, "Total generalized variation," *SIAM Journal on Imaging Sciences*, vol. 3, no. 3, pp. 492-526, 2010.
- [64] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *Journal of Mathematical Imaging and Vision*, vol. 40, no. 1, pp. 120-145, 2011.