

A Novel Retinex-Based Fractional-Order Variational Model for Images With Severely Low Light

Zhihao Gu, Fang Li[✉], Faming Fang[✉], and Guixu Zhang[✉]

Abstract—In this paper, we propose a novel Retinex-based fractional-order variational model for severely low-light images. The proposed method is more flexible in controlling the regularization extent than the existing integer-order regularization methods. Specifically, we decompose directly in the image domain and perform the fractional-order gradient total variation regularization on both the reflectance component and the illumination component to get more appropriate estimated results. The merits of the proposed method are as follows: 1) small-magnitude details are maintained in the estimated reflectance. 2) illumination components are effectively removed from the estimated reflectance. 3) the estimated illumination is more likely piecewise smooth. We compare the proposed method with other closely related Retinex-based methods. Experimental results demonstrate the effectiveness of the proposed method.

Index Terms—Retinex, low-light image, fractional-order, variational model, image enhancement, reflectance, illumination.

I. BACKGROUND

IMAGE enhancement is useful in many scientific areas such as atmospheric sciences, biomedicine and computer vision, which ask for images with excellent qualities for their specific tasks [3]. The word “Retinex” comes from “retina” and “cortex”. Land and Mican [9] first introduced the idea of Retinex in 1971 to model the perception of human vision. The purpose of Retinex theory is to decompose an observed image as the product of illumination and reflectance. Fig. 1 demonstrates four severely low-light images. We can barely see anything clearly in these images due to the extremely low contrast. How to increase the contrast while maintaining excellent image quality (more natural, higher contrast, less noisy and less distorted, and so on) is challenging and exciting.

Manuscript received June 30, 2019; revised October 31, 2019 and December 2, 2019; accepted December 3, 2019. Date of publication December 12, 2019; date of current version January 28, 2020. This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant 61731009, Grant 11671002, and Grant 61871185, in part by the Fundamental Research Funds for the Central Universities, and in part by the Science and Technology Commission of Shanghai Municipality under Grant 18dz2271000. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Dong Xu. (*Corresponding author: Fang Li.*)

Z. Gu is with the School of Mathematical Sciences, East China Normal University, Shanghai 200062, China (e-mail: gzhmsy@126.com).

F. Li is with the Shanghai Key Laboratory of PMMP, School of Mathematical Sciences, East China Normal University, Shanghai 200062, China (e-mail: ffi@math.ecnu.edu.cn).

F. Fang and G. Zhang are with the Shanghai Key Laboratory of Multidimensional Information Processing, School of Computer Science and Technology, East China Normal University, Shanghai 200062, China (e-mail: fmfang@cs.ecnu.edu.cn; gxzhang@cs.ecnu.edu.cn).

Digital Object Identifier 10.1109/TIP.2019.2958144

With the mentioned decomposition, correcting the effects of the lighting to enhance these images becomes possible.

As we know, without any priors, decomposing an observed image to recover both illumination and reflectance is a mathematically ill-posed problem [9]. So, various methods with different priors were then proposed to solve this problem. All these methods can be classified into four types: path-based algorithms [9], [10], center/surround algorithms [11], partial differential equations (PDE) based methods [7], [17] and variational methods [5], [6], [8], [12]–[14], [18], [20], [24], [26].

In the fourth classes, almost all the existing methods perform priors on the integer-order gradient of the illumination and reflectance. Kimmel *et al.* [8] proposed a variational framework to estimate the illumination. They assumed that the illumination was spatially smooth and performed L_2 norm on the gradient of the illumination. Ma *et al.* [14] proposed a total variation (TV) and nonlocal TV regularized model to estimate the reflectance. Following the assumption that the reflectance corresponds to the sharp details and the illumination is spatially smooth, they performed L_1 norm on the gradient of the reflectance and L_2 norm on the difference between the reflectance and the input. Different from the previous methods, Ng *et al.* [18] considered and studied two crucial elements for Retinex to estimate the illumination and the reflectance simultaneously. They assumed spatial smoothness of the illumination and piecewise continuity of the reflectance. In their model, they employed the total variation of the reflectance and performed L_2 norm on the gradient of the illumination. Because of the log-transform, the enhanced results are always over-enhanced. Wang *et al.* [24] borrowed the idea of barrier methods and performed L_2 norm both on the gradient of the illumination and the reflectance, along with some constrained conditions. Based on the analysis of the existing works, Liang *et al.* [13] proposed a high order total variation L_1 decomposition model to simultaneous estimate the illumination and the reflectance. In their model, they perform a total variation on the reflectance and, differently, L_1 norm on the gradient of order 2 of the illumination. Fu *et al.* [5] showed that the log-transform was not ideal for the decomposition and, based on the model of [24], proposed a new weighted gradient based variational model. Park *et al.* [20] proposed a low-light image enhancement method performing L_2 norm on the gradient of the illumination and L_1 norm on the reflectance in the image domain. However, in their model, another term is needed that the illumination should be close to an estimated illumination obtained by performing gamma



Fig. 1. Four images with severely low light.

correction on the given image. Ying *et al.* [26] proposed a weighted L_2 norm model to estimate the exposure ratio in the low-light image and then used their camera response model to adjust the exposure ratio. Li *et al.* [12] proposed a structure-revealing low-light image enhancement model to estimate the illumination, the reflectance, and the noise simultaneously. They performed L_1 norm on the gradient of the illumination, F -norm on both the noise and the difference between the gradient of the reflectance and then adjusted gradient of the input. Gu *et al.* [6] performed L_2 norm on the gradient of the illumination and L_1 norm on the reflectance in the image domain along with a fidelity term to estimate the illumination and the reflectance simultaneously.

Most of the existing variational models are based on the gradient (∇^α , $\alpha \in \{1, 2\}$) whose order is an integer and perform a log-transform on the original image to obtain the addition form from the product form which can reduce the computational complexity of the corresponding algorithm. On the one hand, L_1 norm on the gradient of the reflectance may lead to the piecewise-constant reflectance from the relevant model, which means the removal of small-magnitude details. On the other hand, L_1 norm on the gradient of order 2 (∇^2) leads to piecewise-constant illumination estimated from the corresponding model which may lead to apparent removal of essential components in the corresponding reflectance in [13]. Besides, as shown in [5] and [6], the log-transform may lead to the loss of finer details in the estimated reflectance, which is undesirable. Based on the above observations, in this paper, we explore the effects of a more flexible ‘gradient’ – fractional-order gradient (∇^α , $\alpha \in (1, 2)$) performed both on the illumination and the reflectance for severely low-light images which is seldom explored before. Differing from integer-order gradient, the computation mechanism behind the fractional-order gradient is completely different, that is, the results from this kind of gradient based decomposition models will exhibit another appearances. Besides, we perform the decomposition directly in the image domain. Experimental results and comparisons demonstrate the effectiveness of the proposed model. Our main contributions are: 1) we propose a new constrained and Retinex-based fractional-order variational model for images with severely low light; 2) we prove the existence of a minimizer of the proposed model; 3) we derive an efficient algorithm and give the convergence analysis under some assumptions.

The paper is organized as follows. In section II, we give preliminaries about this paper. In section III, we present the

proposed model along with some theoretical analysis about the existence of the solution. In section IV, we introduce an effective iterative algorithm to solve the proposed model and study the convergence of the algorithm. Then we present some numerical results in section V. In section VI, we discuss the impact of parameters. More extended experiments are given in VII. Finally, concluding remarks are given in section VIII.

II. PRELIMINARIES

In this section, we briefly review the total variation (TV), its fractional-order counterpart and discrete fractional-order gradient operator.

A. Total Variation

For the completeness of this paper, we also give some introductions about TV.

Definition 1 (Total Variation): The total variation of u is

$$TV(u) := \sup_{\phi \in K} \int_{\Omega} (u \cdot \operatorname{div} \phi) dx, \quad (1)$$

where

$K := \{\phi \in (\mathcal{C}_0^\infty(\Omega))^2 \mid \|\phi\|_\infty \leq 1 \text{ for all } x \in \Omega\}$,
 $\operatorname{div} \phi = \sum_{i=1}^d \frac{\partial \phi_i}{\partial x_i}$ and $\frac{\partial \phi_i}{\partial x_i}$ represents the first-order derivative of ϕ_i along x_i direction.

Definition 2 (BV Norm): The BV norm is defined by

$$\|u\|_{BV} = \|u\|_{L^1} + TV(u). \quad (2)$$

Definition 3 (BV Space): The space of functions of bounded variation on Ω is

$$BV(\Omega) = \{u \in L^1(\Omega) \mid TV(u) < +\infty\}. \quad (3)$$

Definition 4 (Weak Topology):* In $BV(\Omega)$, a weak $BV - w^*$ topology is that

$$u_i \xrightarrow[BV-w^*]{*} u \Leftrightarrow u_i \xrightarrow[L^1(\Omega)]{} u \quad (4)$$

and

$$\int_{\Omega} \phi \cdot \nabla u_i dx \longrightarrow \int_{\Omega} \phi \cdot \nabla u dx \quad (5)$$

for all ϕ in $(\mathcal{C}_0^\infty(\Omega))^2$.

Lemma 1 (Lower Semicontinuity): Let $\{u_k(x)\}$ be a sequence from $BV(\Omega)$ converging to a function $u(x)$ in $L^1(\Omega)$, then the following inequality holds.

$$TV(u) \leq \liminf_{k \rightarrow \infty} TV(u^k). \quad (6)$$

Lemma 2 (Compactness): $BV(\Omega)$ is weakly compact with respect to the weak* topology.

B. Fractional-Order Total Variation

Assume that α is a fraction, $\alpha \in R^+$ and a function $h \in L^1([a, b] \in R)$. There are three main definitions of fractional-order derivatives. They are the Riemann-Liouville (RL), the Grünwald-Letnikov (GL), and the Caputo definitions respectively. Our work is mainly based on the RL definition. More details about the other two definitions can be found in [19].

The left and right RL fractional-order derivatives are

$$D_{[a,x]}^\alpha h(x) = \frac{1}{\Gamma(1-\alpha)} \frac{d}{dx} \int_a^x \frac{h(\tau)}{(x-\tau)^\alpha} d\tau, \quad x \in (a, b], \quad (7)$$

and

$$D_{[x,b]}^\alpha h(x) = \frac{-1}{\Gamma(1-\alpha)} \frac{d}{dx} \int_x^b \frac{h(\tau)}{(\tau-x)^\alpha} d\tau, \quad x \in [a, b), \quad (8)$$

where $\Gamma(\cdot)$ represents a gamma function. Based on the above two definitions, the Riesz-RL (central) fractional-order derivative is defined as

$$D_{[a,b]}^\alpha h(x) = \frac{1}{2} (D_{[a,x]}^\alpha h(x) - D_{[x,b]}^\alpha h(x)), \quad x \in (a, b). \quad (9)$$

Referring to [27], we have the following Definitions and Properties:

Definition 5 (Total α -Order Variation): The total α -order variation of u is

$$TV^\alpha(u) := \sup_{\phi \in K} \int_\Omega (-u \operatorname{div}^\alpha \phi) dx, \quad (10)$$

where

$$K := \{\phi \in C_0^1(\Omega, R^d) \mid \|\phi\|_{1,2} \leq 1 \text{ for all } x \in \Omega\},$$

$\operatorname{div}^\alpha \phi = \sum_{i=1}^d \frac{\partial^\alpha \phi_i}{\partial x_i^\alpha}$ and $\frac{\partial^\alpha \phi_i}{\partial x_i^\alpha}$ represents the fractional-order derivative of ϕ_i along x_i direction.

Definition 6 (BV^α Norm): The α - BV norm is defined by

$$\|u\|_{BV^\alpha} = \|u\|_{L^1} + TV^\alpha(u). \quad (11)$$

Definition 7 (BV^α Space): The space of functions of α -bounded variation on Ω is

$$BV^\alpha(\Omega) = \{u \in L^1(\Omega) \mid TV^\alpha(u) < +\infty\}. \quad (12)$$

Definition 8 (W_p^α Norm): For any $p \in N^+$, W_p^α norm is defined as

$$\|u\|_{W_p^\alpha(\Omega)} = \left(\int_\Omega |u|^p dx + \int_\Omega |\nabla^\alpha u|^p dx \right)^{\frac{1}{p}}, \quad (13)$$

where $\nabla^\alpha u = (\frac{\partial^\alpha u}{\partial x_1^\alpha}, \dots, \frac{\partial^\alpha u}{\partial x_n^\alpha})^T$.

Definition 9 (W_p^α Space): The function space embedding with W_p^α norm can be defined as

$$W_p^\alpha(\Omega) = \{u \in L^p(\Omega) \mid \|u\|_{W_p^\alpha(\Omega)} < \infty\}. \quad (14)$$

Definition 10 (weak Topology):* In $BV^\alpha(\Omega)$, a weak $BV^\alpha - w^*$ topology is that

$$u_i \xrightarrow[BV^\alpha-w^*]{*} u \Leftrightarrow u_i \xrightarrow[L^1(\Omega)]{} u \quad (15)$$

and

$$\int_\Omega \phi \cdot \nabla^\alpha u_i dx \longrightarrow \int_\Omega \phi \cdot \nabla^\alpha u dx \quad (16)$$

for all ϕ in $C_0^0(\Omega, R^d)$.

Lemma 3: If $u \in W_1^\alpha(\Omega)$, then $TV^\alpha(u) = \int_\Omega |\nabla^\alpha u| dx$.

Lemma 4 (Lower Semicontinuity): Let $\{u_k(x)\}$ be a sequence from $BV^\alpha(\Omega)$ converging to a function $u(x)$ in $L^1(\Omega)$, then the following inequality holds.

$$TV^\alpha(u) \leq \liminf_{k \rightarrow \infty} TV^\alpha(u^k). \quad (17)$$

Lemma 5 (Compactness): $BV^\alpha(\Omega)$ is weakly compact with respect to the weak* topology.

C. Discrete Fractional-Order Gradient Operator

Assume that $I \in R^{m \times n}$ under Dirichlet boundary condition and $\alpha \in (1, 2)$. According to [21] and [22], with the order α , the discrete fractional-order gradient at point (i, j) in the horizontal and the vertical direction are as follows:

$$\nabla_h^{f(\alpha)} I_{i,j} = \frac{1}{2} \left[\sum_{k=0}^{i+1} c_k^\alpha I_{i-k+1,j} + \sum_{k=0}^{n-i+2} c_k^\alpha I_{i+k-1,j} \right], \quad (18)$$

$$\nabla_v^{f(\alpha)} I_{i,j} = \frac{1}{2} \left[\sum_{k=0}^{j+1} c_k^\alpha I_{i,j-k+1} + \sum_{k=0}^{n-j+2} c_k^\alpha I_{i,j+k-1} \right], \quad (19)$$

where $i, j = 1, \dots, n$ and c_k^α can be computed by the following recursive formulas:

$$c_0^\alpha = 1, \quad c_k^\alpha = \left(1 - \frac{1+\alpha}{k}\right) c_{k-1}^\alpha, \quad k = 1, 2, \dots \quad (20)$$

Set $c = c_0^\alpha + c_2^\alpha$, then we can define the discrete fractional-order gradient of an image I as:

$$\nabla^{f(\alpha)} : R^{m \times n} \rightarrow R^{2m \times n}, \quad \nabla^{f(\alpha)} I := \begin{pmatrix} M_\alpha I \\ I N_\alpha^T \end{pmatrix}, \quad (21)$$

where

$$M_\alpha = \begin{pmatrix} 2c_1^\alpha & c & c_3^\alpha & \cdots & c_m^\alpha \\ c & 2c_1^\alpha & \ddots & \ddots & \vdots \\ c_3^\alpha & \ddots & \ddots & \ddots & c_3^\alpha \\ \vdots & \ddots & \ddots & 2c_1^\alpha & c \\ c_m^\alpha & \cdots & c_3^\alpha & c & 2c_1^\alpha \end{pmatrix} \in R^{m \times m}, \quad (22)$$

$$N_\alpha = \begin{pmatrix} 2c_1^\alpha & c & c_3^\alpha & \cdots & c_n^\alpha \\ c & 2c_1^\alpha & \ddots & \ddots & \vdots \\ c_3^\alpha & \ddots & \ddots & \ddots & c_3^\alpha \\ \vdots & \ddots & \ddots & 2c_1^\alpha & c \\ c_n^\alpha & \cdots & c_3^\alpha & c & 2c_1^\alpha \end{pmatrix} \in R^{n \times n}. \quad (23)$$

Besides, the transposition of the discrete fractional-order gradient of an image I also has its counterpart: $(\nabla^{f(\alpha)})^T I = M_\alpha^T I_h + I_v N_\alpha$.

Specifically, for $\alpha = 1$, according to Eq. (18) - Eq. (20), we can easily get $c_0^\alpha = 1$, $c_1^\alpha = -1$ and $c_i^\alpha = 0$, for

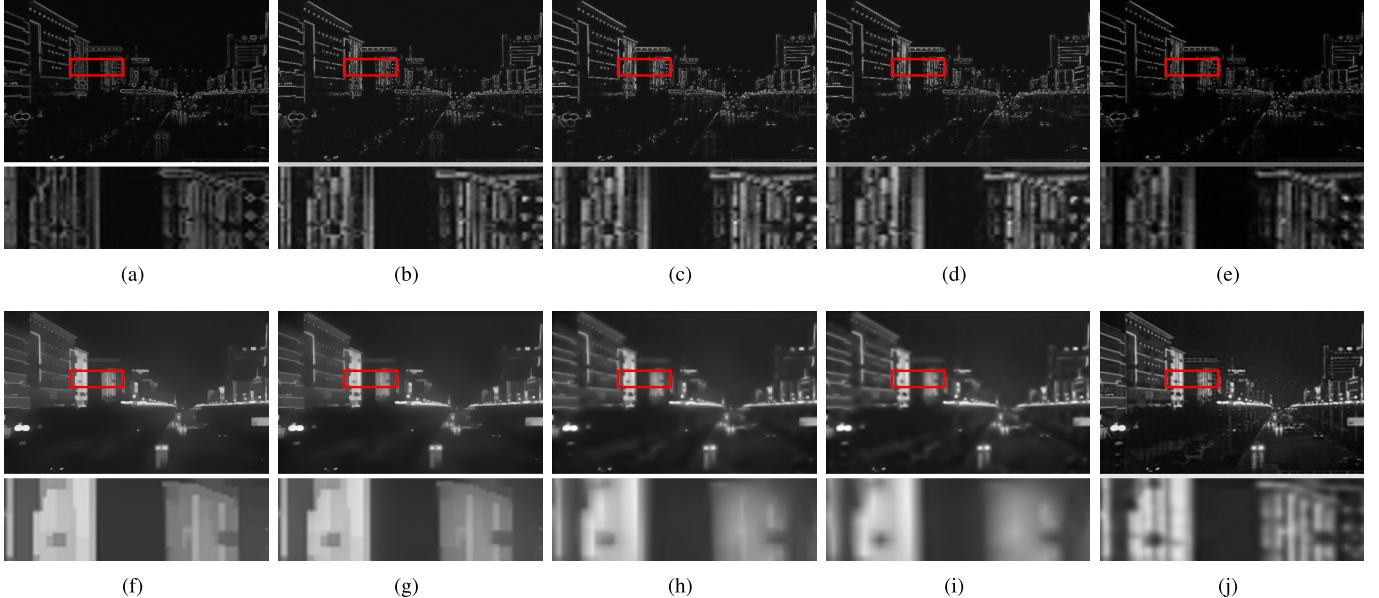


Fig. 2. The gradients, fractional-order gradients of the image I and the corresponding reconstructions (denoted as $\text{ROF}(\cdot)$) from ROF model. (a) ∇I . (b) $\nabla^{1.1}I$. (c) $\nabla^{1.5}I$. (d) $\nabla^{1.9}I$. (e) ∇^2I . (f) $\text{ROF}(\nabla I)$. (g) $\text{ROF}(\nabla^{1.1}I)$. (h) $\text{ROF}(\nabla^{1.5}I)$. (i) $\text{ROF}(\nabla^{1.9}I)$. (j) $\text{ROF}(\nabla^2I)$.

$i = 2, \dots$. Then the discrete fractional-order gradient at point (i, j) in the horizontal and the vertical direction are

$$\nabla_h^{f(1)} I_{i,j} = \frac{1}{2} [I_{i+1,j} + I_{i-1,j} - 2I_{i,j}], \quad (24)$$

$$\nabla_v^{f(1)} I_{i,j} = \frac{1}{2} [I_{i,j+1} + I_{i,j-1} - 2I_{i,j}], \quad (25)$$

while the integral-order gradient at point (i, j) in the horizontal and the vertical direction are

$$\nabla_h I_{i,j} = I_{i+1,j} - I_{i,j}, \quad (26)$$

$$\nabla_v I_{i,j} = I_{i,j+1} - I_{i,j}. \quad (27)$$

Combined with Eq. (24) and Eq. (25), we obtain the following relationship between $\nabla_h^{f(1)}$ and ∇ at point (i, j)

$$\nabla_h^{f(1)} I_{i,j} = \frac{1}{2} [\nabla_h I_{i,j} - \nabla_h I_{i-1,j}], \quad (28)$$

$$\nabla_v^{f(1)} I_{i,j} = \frac{1}{2} [\nabla_v I_{i,j} - \nabla_v I_{i,j-1}]. \quad (29)$$

From the above equations, we can find that 1) $\nabla^{f(1)}$ provides more global information at point (i, j) along the corresponding direction than ∇ ; 2) ∇^2 contains ∇_{hh}^2 , ∇_{vv}^2 , ∇_{hv}^2 and ∇_{vh}^2 , each asking for the gradient of the fixed-direction gradient which indicates that the magnitude of a given image is decided once choosing to use the second-order gradient while the fractional-order is not the case. Besides, we can also observe that the computation mechanism behind the fractional-order gradient is different from that of the integer-order gradient. Because of the non-compatibility of them, we can not obtain the first (or second)-order gradient just setting α be 1 (or 2). These are the most important differences compared to other Retinex-based methods, especially integer-order-gradient-based methods.

Based on the above analysis, we can conclude that $\nabla^{f(\alpha)}$ plays a more flexible role in regulating the gradient. So we have reasons to believe $\nabla^{f(\alpha)}$ may perform better in such

a composition task than the integral-order gradient. In the following sections, for convenience, we denote ∇^α as the discrete fractional-order gradient operator.

III. MODEL AND MINIMIZER

In this section, we first give the motivations of this paper and then present the proposed model. Finally, we give some theoretical analysis of the existence of the minimizer.

A. Motivations

The Retinex theory can be formulated as:

$$I(x, y) = R(x, y)L(x, y), \quad (30)$$

at each point (x, y) , where I , R and L represent the observed image, the reflectance and the illumination, respectively.

It is ill-posed to estimate R and L from Eq. (30) simultaneously. In order to avoid this ill-posedness, different models with various priors are proposed. In most models, they first presume the spatial-smoothness of the illumination and the piecewise-constancy of the reflectance. Then perform the log-transform on Eq. (30) to obtain the additive form

$$i = r + l, \quad (31)$$

where $i = \log I$, $r = \log R$ and $l = \log L$. Finally, with different priors, various regularization terms of the gradient of the reflectance and the illumination, combined with a L_2 fidelity term to enforce the similarity between the input and the addition of l and r , are designed to estimate the illumination and the reflectance, such as $\|\nabla r\|_1$ and $\|\nabla^2 l\|_1$. However, as shown in [5] and [6], using log-transform as preprocessing leads to fine details losing in the estimated reflectance and over-enhanced results.

Secondly, the traditional integral-order gradients (∇ and ∇^2) always lead to strictly piecewise-constant reflectance and piecewise-smooth illumination. Fig. 2 demonstrates the

integral-order gradients, fractional-order gradients of the image I and the reconstructions from the corresponding ROF models. $ROF(\nabla^\alpha I)$ s are reconstructed from

$$\arg \min_u \|\nabla^\alpha u\|_1 + \frac{\lambda}{2} \|u - I\|_2^2. \quad (32)$$

As can be observed, the intensity of $\nabla^\alpha I$ is obviously stronger than ∇I and, with α close to 1, $ROF(\nabla^\alpha I)$ is almost as piecewise constant as ∇I . When increasing α , $\nabla^\alpha I$ seems to be smooth, and the corresponding $ROF(\nabla^\alpha I)$ seems to keep the global boundaries while smooth their inside areas which we still called piecewise smoothness. Therefore, in the sense of changing α , we can flexibly obtain various gradients, which can not be achieved by other Retinex-based methods, especially integer-order-gradient-based methods.

From the above observations, we perform the decomposition directly in the image domain to reduce the loss of details in the estimated reflectance. To have a flexible control over the regularization extent, we use the fractional-order gradient to constrain both the reflectance and the illumination components.

B. Model

The existing Retinex-based models assume that $0 < R < 1$ for the reflectance and $0 < L < \infty$ for the illumination. As we know, almost no materials can absorb and reflect all the lights reaching their surfaces. So with the constraint (30), it's also reasonable to assume $\tau \leq R \leq 1$ and $I \leq L \leq \frac{I}{\tau}$ for R and L respectively, where $0 < \tau < 1$. To sum up, our assumptions on the illumination and the reflectance are as follows:

- The illumination L is fractional-order piecewise smooth.
- The reflectance R is fractional-order piecewise constant.
- R and L satisfy the constraints: $\tau \leq R \leq 1$ and $I \leq L \leq \frac{I}{\tau}$.

Based on the above assumptions, we propose the following fractional-order variational model:

$$\begin{aligned} \min_{R, L} E(R, L) = & \left\{ \frac{1}{2} \int_{\Omega} |RL - I|^2 dx + \lambda_1 \int_{\Omega} |\nabla^\alpha R| dx \right. \\ & \left. + \lambda_2 \int_{\Omega} |\nabla^\beta L| dx \right\} \\ \text{s.t. } & \tau \leq R \leq 1, \quad I \leq L \leq \frac{I}{\tau}, \end{aligned} \quad (33)$$

where λ_1 and λ_2 are positive regularization parameters and α and β are fractional-order parameters which are contained in interval $(1, 2)$. τ is the lower boundary of the reflectance.

The second and the third terms are fractional-order TV regularization terms which enforce the reflectance to contain small-magnitude details and the illumination to be piecewisesmooth respectively. I represents the observed image.

C. Existence of Minimizer

For convenience, we define the following set:

$$\Lambda = \left\{ (R, L) \mid (R, L) \in BV^\alpha(\Omega) \times BV^\beta(\Omega), \right. \\ \left. \tau \leq R \leq 1, I \leq L \leq \frac{I}{\tau} \right\}$$

Then the proposed model (33) can be rewritten as:

$$\min_{(R, L) \in \Lambda} E(R, L). \quad (34)$$

Theorem 1: Let $I \in L^\infty$, then the problem (34) has at least one solution in Λ .

Proof: If R and L are constants, then the energy $E(R, L)$ will be finite and the problem (34) is proper.

Suppose $(R_i, L_i)_{i=1}^\infty \subset \Lambda$ is a minimizing sequence of the problem (34). Then there exists a constant $T > 0$ such that:

$$E(R_i, L_i) \leq T. \quad (35)$$

That is,

$$\frac{1}{2} \int_{\Omega} |R_i L_i - I|^2 dx + \lambda_1 \int_{\Omega} |\nabla^\alpha R_i| dx + \lambda_2 \int_{\Omega} |\nabla^\beta L_i| dx \leq T. \quad (36)$$

which implies

$$\int_{\Omega} |\nabla^\alpha R_i| dx \leq \frac{T}{\lambda_1}. \quad (37)$$

then $\int_{\Omega} |\nabla^\alpha R_i| dx$ is uniformly bounded. Note that

$$\tau \leq R_i \leq 1. \quad (38)$$

So $\{R_i\}$ is uniformly bounded in $L^1(\Omega)$ and $L^2(\Omega)$. Combining the above two results and Lemma 3, we get that $\{R_i\} \in W_1^\alpha$ and then $TV^\alpha(R_i) = \int_{\Omega} |\nabla^\alpha R_i| dx$. Further, $\{R_i\}$ is uniformly bounded in $BV^\alpha(\Omega)$. Thus, based on Lemma 5, there exists a subsequence, also denoted as $\{R_i\}$, converging to $R_* \in BV^\alpha(\Omega)$ such that

$$R_i \xrightarrow[L^1(\Omega)]{} R_* \quad \text{and} \quad R_i \rightharpoonup R_* \in L^2(\Omega). \quad (39)$$

On the other hand, note that

$$I \leq L_i \leq \frac{I}{\tau}, \quad \int_{\Omega} |\nabla^\beta L_i| dx \leq \frac{T}{\lambda_2} \quad (40)$$

we can deduce that $\{L_i\} \in L^\infty$. For the similar reasons, there exists a subsequence converging to $L_* \in BV^\beta(\Omega)$ such that

$$L_i \xrightarrow[L^1(\Omega)]{} L_* \quad \text{and} \quad L_i \rightharpoonup L_* \in L^2(\Omega). \quad (41)$$

Since $R_i L_i \rightharpoonup R_* L_*$ in $L^2(\Omega)$, the lower semicontinuity of the $L^2(\Omega)$ norm yields

$$\liminf_{i \rightarrow \infty} \frac{1}{2} \int_{\Omega} |R_i \cdot L_i - I|^2 dx \geq \frac{1}{2} \int_{\Omega} |R_* \cdot L_* - I|^2 dx. \quad (42)$$

According to Lemma 4, we have

$$\liminf_{i \rightarrow \infty} \lambda_1 \int_{\Omega} |\nabla^\alpha R_i| \geq \lambda_1 \int_{\Omega} |\nabla^\alpha R_*|, \quad (43)$$

and

$$\liminf_{i \rightarrow \infty} \lambda_2 \int_{\Omega} |\nabla^\beta L_i| \geq \lambda_2 \int_{\Omega} |\nabla^\beta L_*|. \quad (44)$$

By summing up the inequalities (42)-(44) together, we obtain

$$\min_{(R, L) \in \Lambda} E(R, L) = \liminf_{i \rightarrow \infty} E(R_i, L_i) \geq E(R_*, L_*). \quad (45)$$

Meanwhile, we have $\tau \leq R_* \leq 1$ and $I \leq L_* \leq \frac{I}{\tau}$. This completes the proof. \square

IV. ALGORITHM AND CONVERGENCE

In this section, we first drive an efficient algorithm to solve the proposed model, and then we analyze the convergence property of the algorithm under some assumptions.

A. Algorithm

In discrete case, the proposed model (33) can be written as:

$$\min_{\{R, L\} \in \Lambda} \frac{1}{2} \|RL - I\|_2^2 + \lambda_1 \|\nabla^\alpha R\|_1 + \lambda_2 \|\nabla^\beta L\|_1. \quad (46)$$

To solve the proposed model efficiently, we adopt the ADMM method [2] to derive the algorithm.

To facilitate an efficient use of alternating minimization, we first introduce four auxiliary variables d_1, u, d_2 and v , and consider the following equivalent model:

$$\begin{aligned} \min_X \quad & \frac{1}{2} \|RL - I\|_2^2 + \lambda_1 \|d_1\|_1 + \lambda_2 \|d_2\|_1 + N_1(R) + N_2(L), \\ \text{s.t. } & u = R, \quad d_1 = \nabla^\alpha u, \quad v = L, \quad d_2 = \nabla^\beta v, \end{aligned} \quad (47)$$

where $X = \{R, L, u, v, d_1, d_2\}$, $N_1(R)$ and $N_2(L)$ denote the indicator functions on the constraint sets $S_1 := \{R | \tau \leq R \leq 1\}$ and $S_2 = \{L | I \leq L \leq \frac{L}{\tau}\}$ respectively. The augmented Lagrangian function of (47) is given by

$$\begin{aligned} \mathcal{L}_A(R, L, u, v, d_1, d_2, \Lambda_u, \Lambda_v, \Lambda_{d_1}, \Lambda_{d_2}) \\ = \frac{1}{2} \|RL - I\|_2^2 + \lambda_1 \|d_1\|_1 + \lambda_2 \|d_2\|_1 + N_1(R) + N_2(L) \\ + \langle \Lambda_u, u - R \rangle + \langle \Lambda_v, v - L \rangle + \langle \Lambda_{d_1}, d_1 - \nabla^\alpha u \rangle \\ + \langle \Lambda_{d_2}, d_2 - \nabla^\beta v \rangle + \frac{\sigma_1}{2} \|u - R\|_2^2 + \frac{\sigma_2}{2} \|v - L\|_2^2 \\ + \frac{\sigma_3}{2} \|d_1 - \nabla^\alpha u\|_2^2 + \frac{\sigma_4}{2} \|d_2 - \nabla^\beta v\|_2^2, \end{aligned} \quad (48)$$

where $\Lambda_u, \Lambda_{d_1}, \Lambda_v, \Lambda_{d_2}$ are Lagrangian multipliers, $\sigma_i \geq 0$, $i = 1, \dots, 4$ are penalty parameters. The ADMM algorithm for (47) is derived by minimizing \mathcal{L}_A with respect to R, L, u, v, d_1, d_2 , one at a time while fixing others at their most recent values, i.e.,

$$\begin{aligned} R^{k+1} &= \arg \min_R \mathcal{L}_A(R, L^k, u^k, v^k, d_1^k, d_2^k, \Lambda_u^k, \Lambda_v^k, \\ &\quad \Lambda_{d_1}^k, \Lambda_{d_2}^k), \\ L^{k+1} &= \arg \min_L \mathcal{L}_A(R^{k+1}, L, u^k, v^k, d_1^k, d_2^k, \Lambda_u^k, \Lambda_v^k, \\ &\quad \Lambda_{d_1}^k, \Lambda_{d_2}^k), \\ u^{k+1} &= \arg \min_u \mathcal{L}_A(R^{k+1}, L^{k+1}, u, v^k, d_1^k, d_2^k, \Lambda_u^k, \Lambda_v^k, \\ &\quad \Lambda_{d_1}^k, \Lambda_{d_2}^k), \\ v^{k+1} &= \arg \min_v \mathcal{L}_A(R^{k+1}, L^{k+1}, u^{k+1}, v, d_1^k, d_2^k, \Lambda_u^k, \Lambda_v^k, \\ &\quad \Lambda_{d_1}^k, \Lambda_{d_2}^k), \\ d_1^{k+1} &= \arg \min_d \mathcal{L}_A(R^{k+1}, L^{k+1}, u^{k+1}, v^{k+1}, d_1, d_2^k, \\ &\quad \Lambda_u^k, \Lambda_v^k, \Lambda_{d_1}^k, \Lambda_{d_2}^k), \\ d_2^{k+1} &= \arg \min_q \mathcal{L}_A(R^{k+1}, L^{k+1}, u^{k+1}, v^{k+1}, d_1^k, \\ &\quad d_2, \Lambda_u^k, \Lambda_v^k, \Lambda_{d_1}^k, \Lambda_{d_2}^k). \end{aligned} \quad (49)$$

and then updating the Lagrangian multipliers $\Lambda_u, \Lambda_v, \Lambda_{d_1}, \Lambda_{d_2}$. It is easy to derive that the first-order optimal condition of the subproblems in (49) are

$$\begin{aligned} 0 &\in L^k(R^{k+1}L - I) - \Lambda_u^k + \sigma_1(R^{k+1} - u^k) \\ &\quad + \partial N_1(R^{k+1}), \end{aligned} \quad (50a)$$

$$\begin{aligned} 0 &\in R^{k+1}(L^{k+1} - I) - \Lambda_v^k + \sigma_2(L^{k+1} - v^k) \\ &\quad + \partial N_2(L^{k+1}), \end{aligned} \quad (50b)$$

$$\begin{aligned} 0 &= \Lambda_u^k - \nabla^{\alpha T} \Lambda_{d_1}^k + \sigma_1(u^{k+1} - R^{k+1}) + \sigma_3 \nabla^{\alpha T} \\ &\quad \times (\nabla^\alpha u^{k+1} - d_1^k), \end{aligned} \quad (50c)$$

$$\begin{aligned} 0 &= \Lambda_v^k - \nabla^{\beta T} \Lambda_{d_2}^k + \sigma_2(v^{k+1} - L^{k+1}) + \sigma_4 \nabla^{\beta T} \\ &\quad \times (\nabla^\beta v^{k+1} - d_2^k), \end{aligned} \quad (50d)$$

$$0 \in \lambda_1 \partial \|d_1^{k+1}\|_1 + \Lambda_{d_1}^k + \sigma_3(d_1^{k+1} - \nabla^\alpha u^{k+1}), \quad (50e)$$

$$0 \in \lambda_2 \partial \|d_2^{k+1}\|_1 + \Lambda_{d_2}^k + \sigma_4(d_2^{k+1} - \nabla^\beta v^{k+1}). \quad (50f)$$

In the following, we solve equations (50a)-(50f) one by one. Since the constraint set S_1 of R in N_1 is convex, from (50a), it is direct to derive the updating formula of R as

$$R^{k+1} = \mathcal{P}_1 \left(\frac{L^k I + \Lambda_u^k + \sigma_1 u^k}{(L^k)^2 + \sigma_1} \right), \quad (51)$$

where \mathcal{P}_1 denotes the projection on the convex set S_1 which is defined as $(\mathcal{P}_1(z))_{ij} = \min(\max(z_{ij}, \tau), 1)$. Similarly, since the constraint set S_2 of L in N_2 is also convex, from (50b), we can derive the updating formula of L as

$$L^{k+1} = \mathcal{P}_2 \left(\frac{R^{k+1} I + \Lambda_v^k + \sigma_2 v^k}{(R^{k+1})^2 + \sigma_2} \right), \quad (52)$$

where \mathcal{P}_2 denotes the projection on the convex set C_2 which is defined as $(\mathcal{P}_2(z))_{ij} = \min(\max(z_{ij}, I_{ij}), \frac{I_{ij}}{\tau_1})$.

From (50c) and (50d), we get the following formulas of u and v as

$$\sigma_3 \nabla^{\alpha T} \nabla^\alpha u^{k+1} + \sigma_1 u^{k+1} = C_1, \quad (53)$$

$$\sigma_4 \nabla^{\beta T} \nabla^\beta v^{k+1} + \sigma_2 v^{k+1} = C_2. \quad (54)$$

where

$$C_1 = \nabla^{\alpha T} \left(\Lambda_{d_1}^k + \sigma_3 d_1^k \right) - \Lambda_u^k + \sigma_1 R^{k+1},$$

$$C_2 = \nabla^{\beta T} \left(\Lambda_{d_2}^k + \sigma_4 d_2^k \right) - \Lambda_v^k + \sigma_2 L^{k+1}.$$

Then by simple operations, we transform the formulas above into the following forms

$$A_1 u^{k+1} + u^{k+1} B_1 = C_1, \quad (55)$$

$$A_2 v^{k+1} + v^{k+1} B_2 = C_2. \quad (56)$$

where A_1, A_2, B_1 , and B_2 are defined as

$$A_1 = \left(\sigma_3 M_\alpha^T M_\alpha + \sigma_1 E \right),$$

$$A_2 = \left(\sigma_4 M_\beta^T M_\beta + \sigma_2 E \right),$$

$$B_1 = \sigma_3 N_\alpha^T N_\alpha,$$

$$B_2 = \sigma_4 N_\beta^T N_\beta,$$

The updating formulas of u and v can be given by

$$u^{k+1} = \text{Sylvester}(A_1, B_1, C_1), \quad (57)$$

$$v^{k+1} = \text{Sylvester}(A_2, B_2, C_2), \quad (58)$$

Algorithm 1 The Proposed Retinex Algorithm

- Input I , $maxiter > 0$, and $tol > 0$. Set $\lambda_1 > 0$, $\lambda_2 > 0$, $\sigma_1 > 0$, $\sigma_2 > 0$, $\sigma_3 > 0$, $\sigma_4 > 0$.
- Initialize $L^0 = I$, $v^0 = I$, $q^0 = \nabla I$ and $R^0, u^0, d^0, \Lambda_u^0, \Lambda_v^0, \Lambda_d^0, \Lambda_q^0$ are zero matrices with appropriate sizes.
- **for** $k = 1, \dots, maxiter$ **do**

 - Update R^k by (51);
 - Update L^k by (52);
 - Update u^k by (57);
 - Update v^k by (58);
 - Update d_1^k by (60);
 - Update d_2^k by (61);
 - Update Λ_u^k by (63);
 - Update Λ_v^k by (64);
 - Update $\Lambda_{d_1}^k$ by (65);
 - Update $\Lambda_{d_2}^k$ by (66);
 - if** a stopping criterion is reached
 then exit and output (R^k, L^k)
 end if

- **end for**

where $X = \text{Sylvester}(A, B, C)$ denotes the operation solving the Sylvester-equation

$$AX + XB = C. \quad (59)$$

and $A_1 = (\sigma_3 M_\alpha^T M_\alpha + \sigma_1 E)$, $B_1 = \sigma_3 N_\alpha^T N_\alpha$, $A_2 = (\sigma_4 M_\beta^T M_\beta + \sigma_2 E)$, and $B_2 = \sigma_4 N_\beta^T N_\beta$. Finally, from (50e) and (50f), by using the definition of sub-gradient of L_1 norm, we can derive that the updating formulas of d_1 and d_2 are given by soft shrinkage:

$$d_1^{k+1} = \mathcal{S}(\nabla^\alpha u^{k+1} - \frac{\Lambda_{d_1}^k}{\sigma_3}, \frac{\lambda_1}{\sigma_3}), \quad (60)$$

$$d_2^{k+1} = \mathcal{S}(\nabla^\beta v^{k+1} - \frac{\Lambda_{d_2}^k}{\sigma_4}, \frac{\lambda_2}{\sigma_4}), \quad (61)$$

where \mathcal{S} denotes the soft shrinkage operator defined as

$$(\mathcal{S}(s, c))_{ij} := \frac{s_{ij}}{|s_{ij}|} * \max(|s_{ij}| - c, 0). \quad (62)$$

According to the standard ADMM algorithm, the updating formulas for the Lagrangian multipliers $\Lambda_u, \Lambda_v, \Lambda_d, \Lambda_q$ are given by

$$\Lambda_u^{k+1} = \Lambda_u^k + \sigma_1(u^{k+1} - R^{k+1}), \quad (63)$$

$$\Lambda_v^{k+1} = \Lambda_v^k + \sigma_2(v^{k+1} - L^{k+1}), \quad (64)$$

$$\Lambda_{d_1}^{k+1} = \Lambda_{d_1}^k + \sigma_3(d_1^{k+1} - \nabla^\alpha u^{k+1}), \quad (65)$$

$$\Lambda_{d_2}^{k+1} = \Lambda_{d_2}^k + \sigma_4(d_2^{k+1} - \nabla^\beta v^{k+1}). \quad (66)$$

Finally, the proposed numerical algorithm for solving model (46) can be summarised in Algorithm 1. The stopping criterion is the following relative errors between the successive iterations satisfies

$$\epsilon_R = ||R^k - R^{k-1}|| / ||R^{k-1}|| < tol,$$

or

$$\epsilon_L = ||L^k - L^{k-1}|| / ||L^{k-1}|| < tol.$$

Considering the computational complexity of Algorithm 1, the most costly steps are the iterations of u^k and v^k , with the order of complexity $O(m^3 + n^3)$, because of the need to solve the Sylvester-equation. Note that the numerical solution of the Sylvester equation we used here is the famous Bartels-Stewart algorithm, which consists of computing the real Schur decompositions by a QR algorithm, and solving the resulting triangular system via back-substitution [1]. Besides, as the complexity of both the projection and soft shrinkage operators are $O(mn)$, all of the other iterative formulas (i.e., (51), (52), (60), (61) and (63)-(66)) have computational complexity $O(mn)$. Thus the overall order of complexity per iteration of the Algorithm 1 is $O(m^3 + n^3)$.

B. Convergence

It is difficult to prove the global convergence of the proposed model (46) due to its non-convexity. Inspired by the work in [25] analyzing the convergence of the alternating direction algorithm for nonnegative matrix completion, we give the convergence property of the proposed Algorithm 1 under some assumptions.

For convenience, we introduce two notations: $X = (R, L, u, v, d_1, d_2)$ and $Y = (\Lambda_u, \Lambda_v, \Lambda_{d_1}, \Lambda_{d_2})$. We assume the sequence (X^k, Y^k) is generated by Algorithm 1. Then we give the convergence of the sequence.

Lemma 6: The point (X, Y) is a KKT point of the problem (46) if it satisfies the following formulas:

$$\begin{aligned} 0 &\in RL^2 - LI - \Lambda_u + \partial N_1(R), \\ 0 &\in LR^2 - RI - \Lambda_v + \partial N_2(L), \\ 0 &= \Lambda_u - (\nabla^\alpha)^T \Lambda_{d_1}, \\ 0 &= \Lambda_v - (\nabla^\beta)^T \Lambda_{d_2}, \\ 0 &\in \lambda_1 \partial \|d_1\|_1 + \Lambda_{d_1}, \\ 0 &\in \lambda_2 \partial \|d_2\|_1 + \Lambda_{d_2}, \\ u &= R, \\ v &= L, \\ d_1 &= \nabla^\alpha u, \\ d_2 &= \nabla^\beta v. \end{aligned} \quad (67)$$

Proof: Following the definition of KKT point for an optimal problem, one can easily obtain the above formulas. \square

Theorem 2 (Convergence): If the sequence $\{X^k\}$ and $\{Y^k\}$ satisfy

$$X^{k+1} - X^k \rightarrow 0, \quad Y^{k+1} - Y^k \rightarrow 0 \quad (68)$$

as $k \rightarrow \infty$, then any accumulation point of $\{(R^k, L^k)\}$ satisfies the KKT condition for problem (46).

The proof can be found in the appendix.

V. NUMERICAL RESULTS**A. Experimental Settings**

All the experiments are performed on a PC with Intel Core i7 CPU 6800k, 64GB RAM with MATLAB to be the programming language. We compare the proposed model (46)

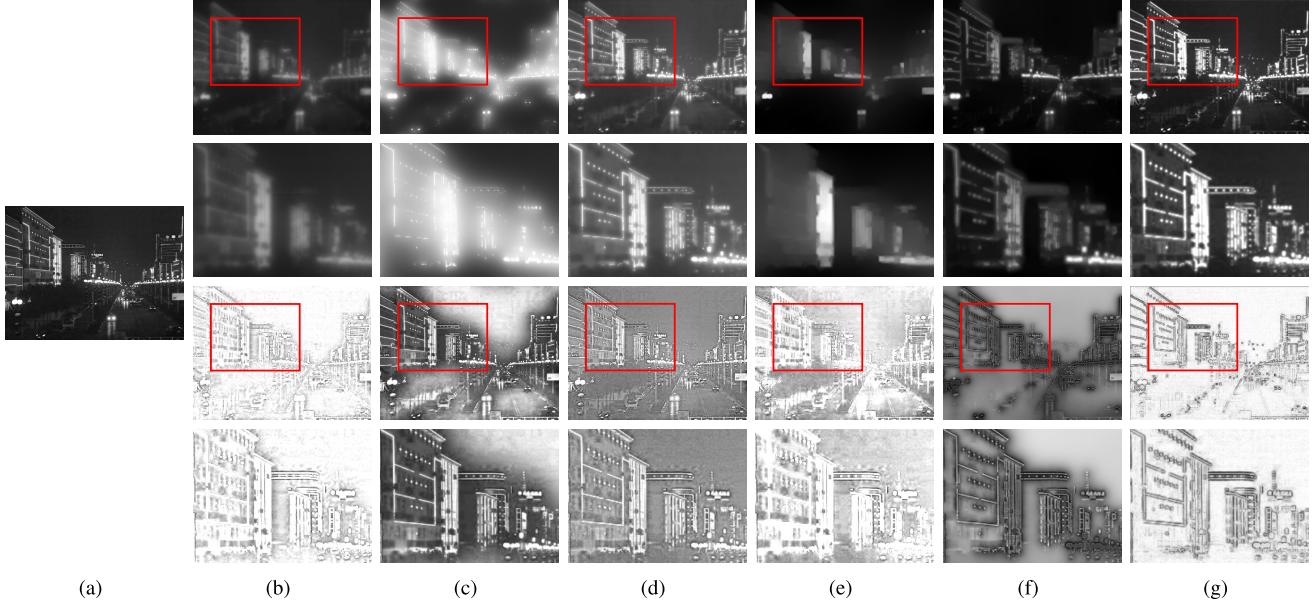


Fig. 3. The decomposition results of different methods. The first and third rows are the estimated illumination and reflection respectively. The second and the fourth rows are the corresponding magnified red rectangle areas. (a) Input. (b) TVMR [18]. (c) WVM [5]. (d) LIME [20]. (e) CRM [26]. (f) SRLLE [12]. (g) the proposed.

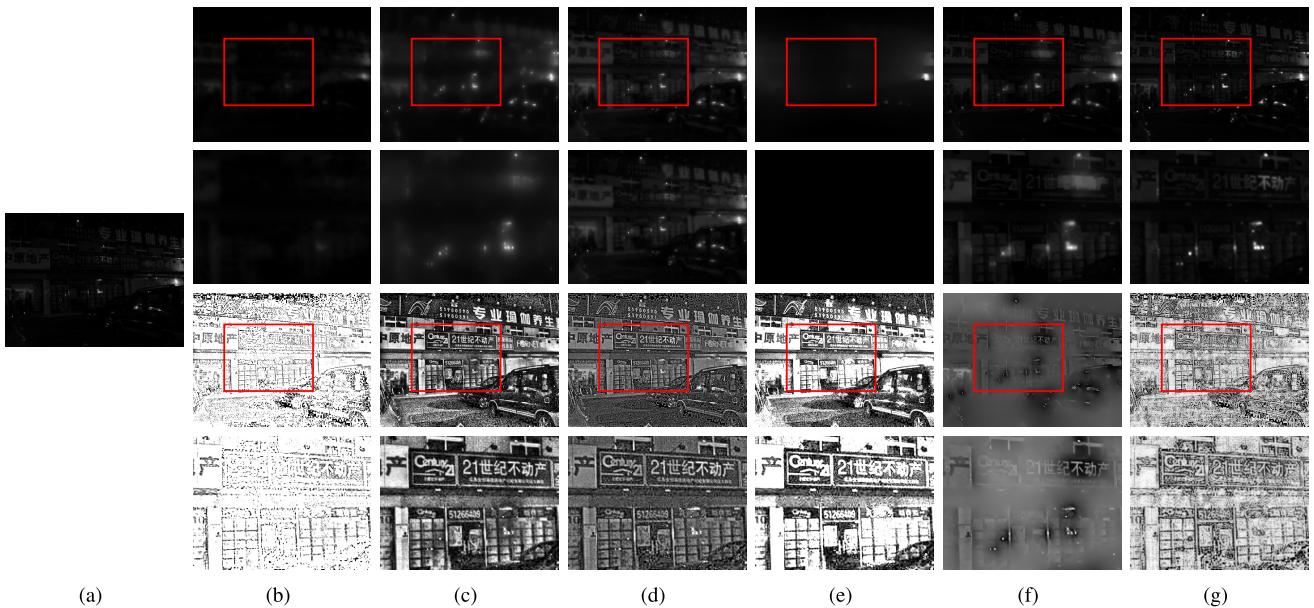


Fig. 4. The decomposition results of different methods. The first and third rows are the estimated illumination and reflection respectively. The second and the fourth rows are the corresponding magnified red rectangle areas. (a) Input. (b) TVMR [18]. (c) WVM [5]. (d) LIME [20]. (e) CRM [26]. (f) SRLLE [12]. (g) the proposed.

with a few state-of-the-art methods, including TVMR [18], WVM [5], LIME [20], CRM [26] and SRLLE [12]. To be fair, we use the codes obtained from authors' websites with default parameters directly. In all the experiments, we empirically set the parameters $\lambda_1 = 0.01$, $\lambda_2 = 0.1$, $\sigma_1 = 0.02$, $\sigma_2 = 10$, $\sigma_3 = 0.02$, $\sigma_4 = 10$, $\alpha = 1.3$ and $\beta = 1.7$ respectively. In subsection V-C, we set $\gamma_1 = 4.6$ and $\gamma_2 = 2.6$ respectively. Besides, taking the convergence speed and computation time into consideration, we set the stop parameters *maxiter* as 25 and the lower bound of reflectance τ as 10^{-4} . Following the previous methods, for color images, we first transform them into the HSV space and then apply the proposed model to

decompose the V-channel as the reflectance and the illumination. For image enhancement, we perform gamma correction on both the estimated reflectance and the illumination and multiply them to get the corrected results. After that, turning it back into the RGB space to obtain the final results (see Fig. 5).

B. Decomposition Results

Fig. 3 and Fig. 4 demonstrate the decomposition results from different methods performed on Fig. 1(b) and Fig. 1(c), respectively. Red rectangle areas are to be compared. As is mentioned in [18], the estimated reflectance images from

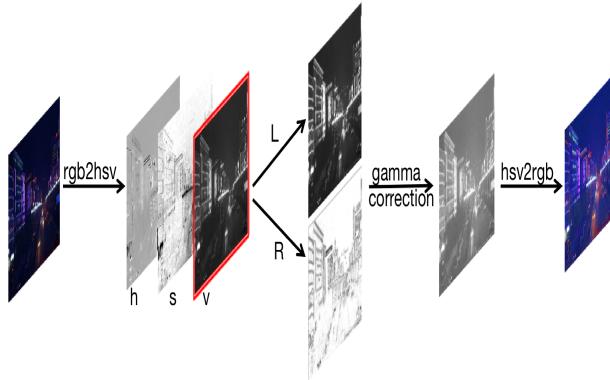


Fig. 5. RGB image enhancement flow chart.

TVMR tend to be over-enhanced, which means loss of fine details and can be obviously observed in both Fig. 3 (b) and Fig. 4 (b) (wrong structures of the building and the vague words respectively) while our results contain proper structures and some details. WVM can obtain reflectance with more information. However, the estimated reflectance not only preserves details but also illumination components such as some dark low-light areas which make the estimated result look low-light. Besides, the estimated illumination images are too spatially smooth that the edges can barely be recognized. In contrast, our reflectance images preserve details and do not contain illumination components, and the estimated illumination images are more piecewise smooth. LIME uses a fidelity term to enforce the similarity between the estimated illumination and the gamma correction version of the initial illumination filtered on the input. So the estimated illumination images are more like the filtered ones. Moreover, the estimated reflectance images still contain the illumination components. CRM introduces the response characteristics of cameras to decompose images. The estimated L seems to be piece-constant and eliminate small light areas. The estimated R contains over-bright areas and hence loses some details. SRLLE considers the noise in their model to suppress its effects in low-light regions. However, for severely low-light images, they are full of low-light areas, and some details may be considered as noise. As a result, the estimated results will be blurred (see the amplified areas in Fig. 3 (f) and Fig. 4 (f)). Though our results may contain some noise which is kept as finer details in the reflectance, they are still comparatively natural.

C. Severely low-Light Image Enhancement

As we know, Retinex is to decompose an observed image into illumination and reflectance. Ideally, once obtaining the ground truth of illumination and reflectance, correcting the estimated illumination can lead to the enhanced result. However, estimating both the illumination and the reflectance is ill-posed. Almost all the Retinex-based image enhancement methods perform gamma correction directly and only on the estimated illumination from their models, and then multiply the corresponding reflectance using the following equation to

achieve the enhanced results.

$$I_{en} = R' \cdot L', \quad (69)$$

where

$$R' = R^1, \quad L' = W\left(\frac{L}{W}\right)^{\frac{1}{\gamma}}, \quad (70)$$

γ is a gamma correction parameter, and W equals to 255 for an 8-bit image.

Actually, they pre-assume that the estimated illumination image from their models is extremely close to the ground truth. Unfortunately, no one can guarantee accuracy. So from this observation, we prefer the more general one as follows

$$I_{en} = R' \cdot L', \quad (71)$$

where

$$R' = R^{\frac{1}{\gamma_0}}, \quad L' = W\left(\frac{L}{W}\right)^{\frac{1}{\gamma}}. \quad (72)$$

Fig. 6 illustrates some enhanced results from different methods on Fig. 1. Results from TVMR is obviously over-enhanced and unnatural. Results from WVM and LIME look alike and still have a comparatively low contrast (see the woman's face and the trees beside the road). Using the response characteristics of cameras makes results from CRM have high contrast and look better; however, there is still color distortion and over-bright of the light. Results from SRLLE have lower contrast compared with others. As discussed in subsection V-B, considering the noise estimation in their model and the low-lightness of the images, SRLLE may perform denoising in dark areas. It may lead to some unpleasant consequences such as blurriness and low contrast (the woman's blurred face and the overall visual effects). Results from the proposed model have comparatively high contrast and look better. Some areas behind darkness such as the woman's face can be observed though there exists some noise.

Besides subjective visual comparisons, we also use three image quality assessment metrics to assess the enhanced results. They are blind image quality assessment metric — natural image quality evaluator (NIQE) [16], reference-based light order error (LOE) [23] and information entropy (IE) respectively. The NIQE is devised to predict the quality of distorted images with as little prior knowledge of the images or their distortions as possible. The lower the NIQE is, the better image quality the result has. For LOE, it is used to assess the naturalness preservation of an image. A lower value represents a better lightness order preserved in the image. On the contrary, larger IE corresponds to the result containing more image information.

Fig. 7 demonstrates NIQE, LOE and IE results from different enhancement methods on Fig. 1 and 100 severely low-light images in Pku-EAQA dataset [4] respectively. From Fig. 7, we can find that the enhanced results from the proposed method have lower averaged NIQE and LOE values. Note that, for some images, results from TVMR have a little higher IE value than ours. We note that TVMR suffers from over-enhancement and may introduce some extra information to enlarge the IE value. Averagely, we have the highest IE value among all. Overall, the enhanced results by the proposed

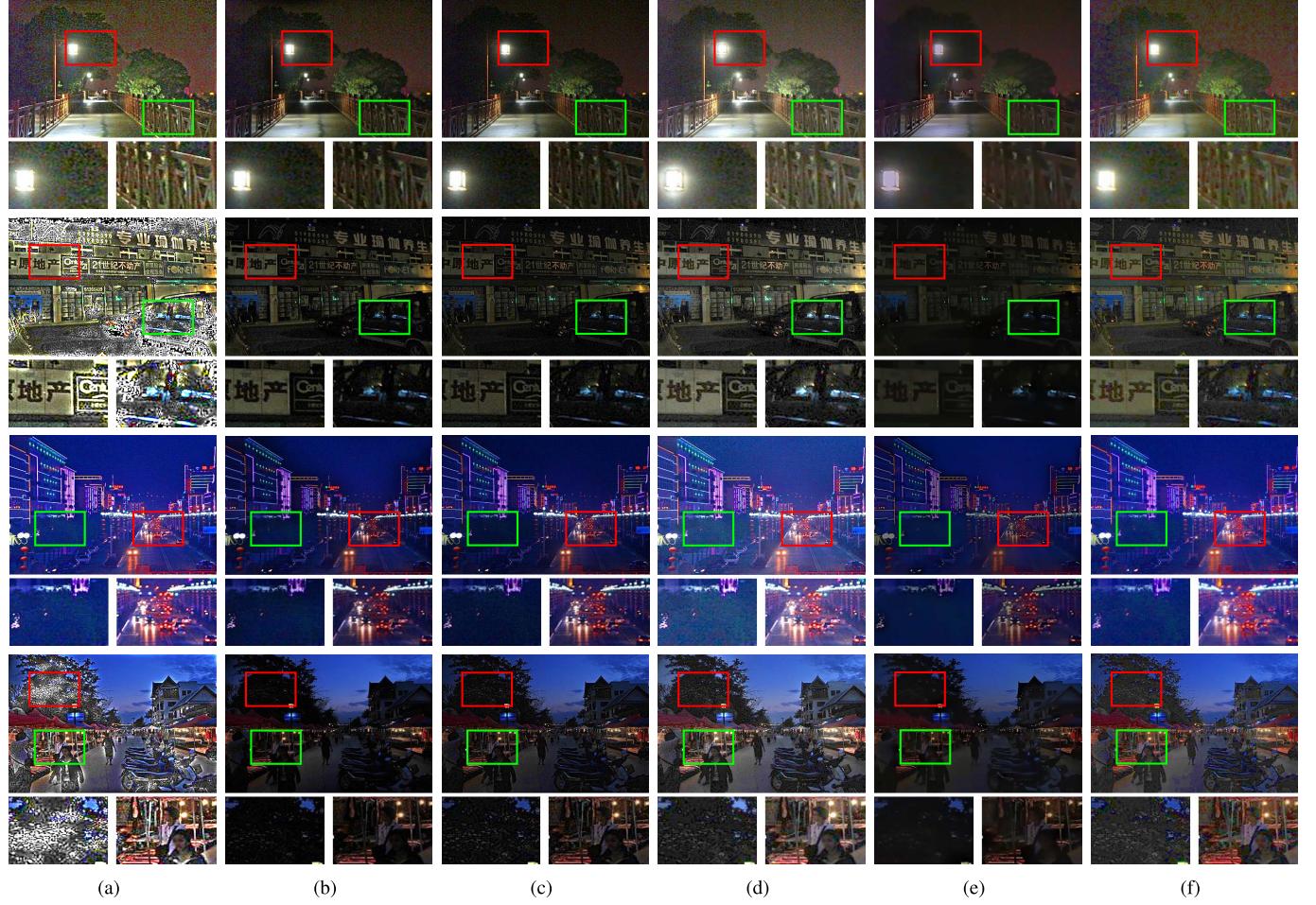


Fig. 6. Enhanced results by different methods on Fig. 1. (a) TVMR [18]. (b) WVM [5]. (c) LIME [20]. (d) CRM [26]. (e) SRLLE [12]. (f) the proposed.

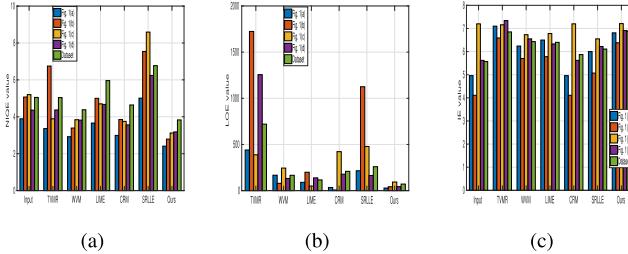


Fig. 7. (a), (b), and (c) are NIQE, LOE, and IE results of different methods on Fig. 1 and 100 severely low-light images, respectively.

method have better qualities and it is better at enhancing images with severely low light than the other compared methods.

D. Noise Sensitivity

We test the noise sensitivity of different methods on synthesized low-light and noisy images. We take the source images with almost normal contrast from Berkeley segmentation dataset (BSD) (<https://www.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/>). The synthesis steps are as follows: performing gamma correction with $\gamma = 2.2$ on the source images first and then add Poisson noise and white Gaussian

noise with $\sigma = 5$ respectively. Structure similarity index (SSIM), peak signal to noise ratio (PSNR) and LOE are used to measure the sensitivity.

Fig. 8 demonstrates the results of different methods on two images from the BSD train set. The first column shows the source images and the second column shows the synthetic low-light and noisy images. As shown in the third column, TVMR leads to severely over-enhanced and distorted results. In the fourth and fifth columns, results from WVM and LIME look alike but still have comparatively low contrast. The results of WVM seem cleaner than LIME. Results from CRM seem to look better, but they suffer from some color distortion and over-bright (the sixth column). The results of SRLLE in the last second column are low-light, noisy and having distortions. However, our results in the last column are cleaner while having a comparatively higher contrast. Overall, when processing noisy images, all these enhancement methods are sensitive to noise more or less.

In TABLE I, we report the quantitative measures of the sensitivity results on 100 synthetic low-light noisy images. The source images are from the BSD test set. Our method and LIME have the highest SSIM values. In term of PSNR, our method is the best and CRM is the second best. In terms of LOE, LIME is the best and our method is the second best.

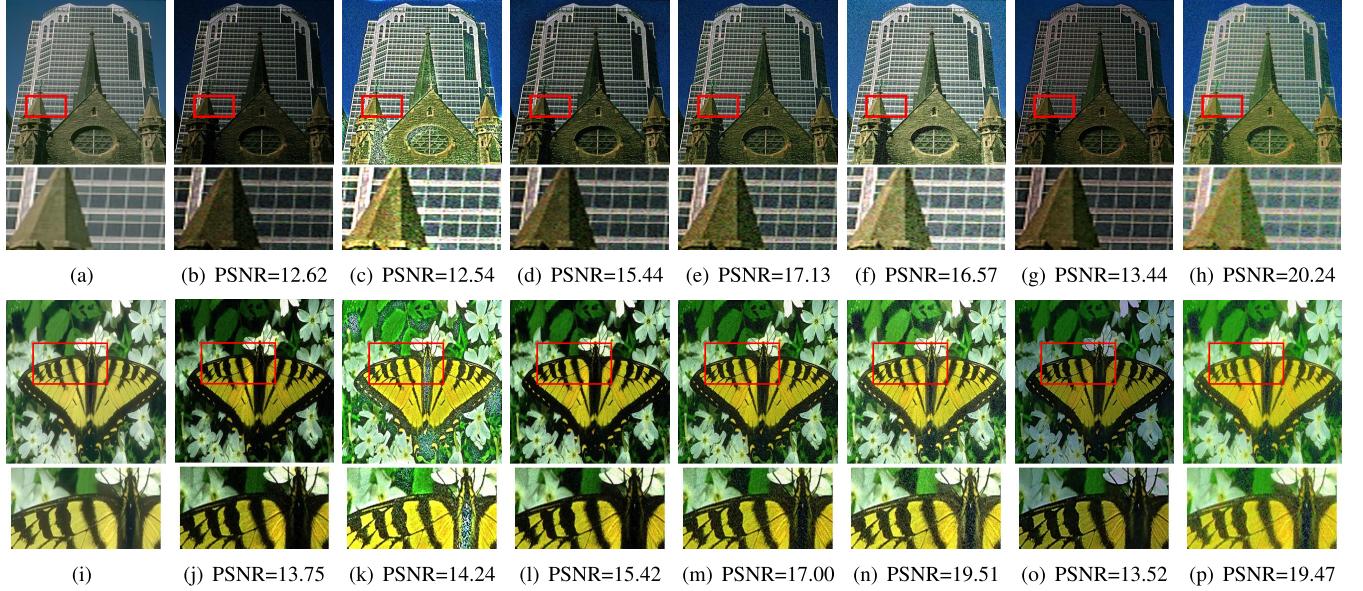


Fig. 8. Examples of noise sensitivity of different methods. (a) and (i) Source images. (b) and (j) Low light and noisy images. (c) and (k) TVMR [18]. (d) and (l) WVM [5]. (e) and (m) LIME [20]. (f) and (n) CRM [26]. (g) and (o) SRLLE [12]. (h) and (p) the proposed. The number under each image is PSNR value (db).

TABLE I
AVERAGE SSIM AND PSNR VALUES ON 100 SYNTHETIC
LOW-LIGHT AND NOISY IMAGES

| | [18] | [5] | [20] | [26] | [12] | Ours |
|------|---------|--------|---------------|--------|---------|--------------|
| SSIM | 0.55 | 0.64 | 0.66 | 0.47 | 0.64 | 0.66 |
| PSNR | 12.90 | 16.02 | 17.03 | 18.35 | 13.88 | 19.84 |
| LOE | 1910.80 | 336.47 | 247.11 | 718.19 | 6912.90 | 322.06 |

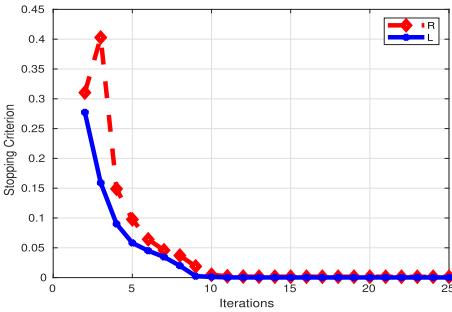


Fig. 9. The relationship between stopping criterion and the iterations.

E. Convergence Speed

In this subsection, we test the convergence speed of the proposed method. Fig. 9 illustrates the convergence curve. In the beginning, due to the initial values of the illumination and the reflectance, there exists a fluctuation. As the iteration increases, the curve descends fast and then tends to become smooth. Intuitively, the proposed algorithm converges within 25 iterations, which implies that the proposed algorithm has a fast convergence speed.

For the 100 severely low-light test images in Fig. 7 (with size 400×300), the average computational time for all the compared methods are listed as follows: TVMR [18] 0.12s, WVM [5] 12.35s, LIME [20] 1.07s, CRM [26] 0.13s,

SRLLE [12] 17.24s, and Ours 7.79s. Our method is faster than WVM and SRLLE but slower than others. As analyzed in subsection IV.A, the most time-consuming part of our method is solving the Sylvester-like equations by MATLAB's build-in subroutine sylvester.m. A faster version can hugely reduce the computational time, which is left as our future work.

VI. IMPACT OF PARAMETERS

In this section, taking Fig 1(b) for example to discuss the impacts of the parameters (λ_1 and λ_2 , α and β , γ_1 and γ_2) on both the decomposition and the enhancement results.

A. Regularization Parameters – λ_1 , λ_2

Fig. 10 demonstrates some visual impacts of regularization parameters pairs (λ_1 , λ_2) on the NIQE value. We list five groups of images corresponding to five regularization parameters pairs, and they are almost no visual differences. Fig. 11 (a) plots the corresponding NIQE, LOE and IE curves where λ_1 ranges in 0.01, 0.1 and 1 and λ_2 is chosen in 0.01, 0.1 and 1. Fig. 11 (b) and (c) demonstrate the NIQE, LOE and IE curves respect to λ_2 and λ_1 with another parameter fixed respectively. When fixing λ_1 , for the LOE curve, it ascends first and then descends while for IE curve, it ascends first and then maintains unchanged. However, for the NIQE curve, it appears the opposite situation that it descends first and ascends later. When fixing λ_2 , the LOE and NIQE curves have the same tendency that they descend first and then ascend. However, IE value keeps ascending. From these curves, it can be observed that (0.01, 0.1) has a better balance between image distortion and naturalness preservation.

B. Fractional Orders – α , β

Fig. 12 demonstrates the impacts of α and β on image quality. From the 3D curve in Fig. 12(a), we can find that

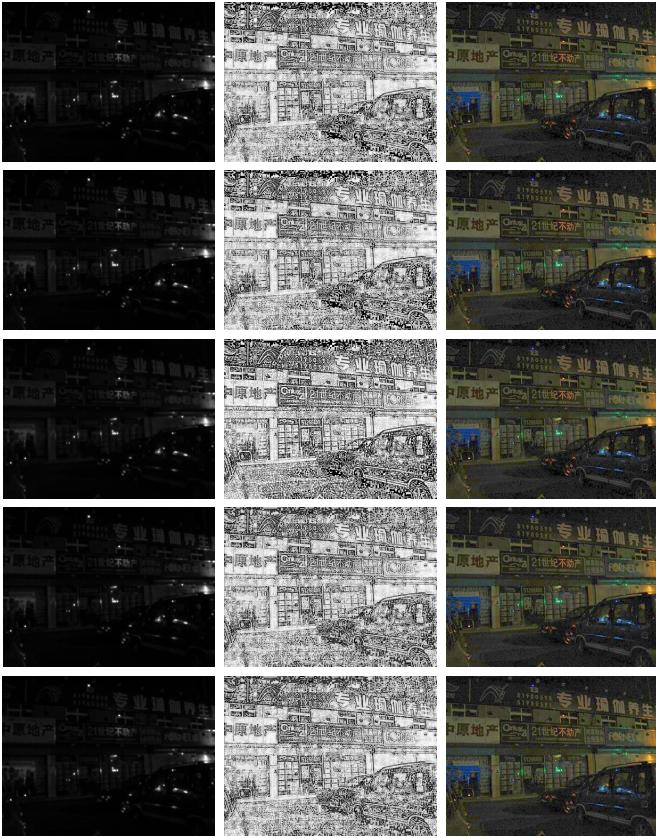


Fig. 10. Examples of impacts of regularization pairs (λ_1, λ_2) . From the top to the bottom: (0.01, 0.01), (0.01, 0.1), (0.01, 1), (0.1, 0.1), (1, 0.1) respectively. From left column to right column: the estimated illumination, the estimated reflectance and the enhanced result.

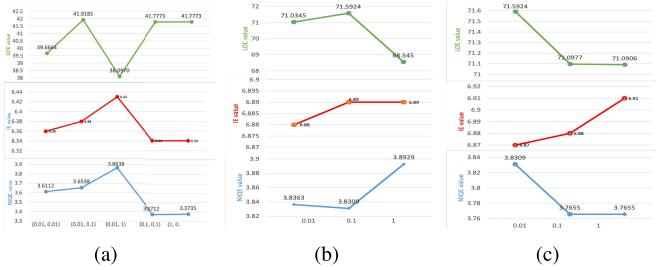


Fig. 11. Impacts of different regularization parameters on average NIQE and LOE results of enhanced Fig 1(b) and 100 severely low-light images respectively. (a) impacts of different (λ_1, λ_2) pairs. (b) fixing $\lambda_1 = 0.01$ and changing λ_2 . (c) fixing $\lambda_2 = 0.01$ and changing λ_1 .

NIQE value is the inverse ratio to β , and for the same β , changes of α have few impacts on it. Besides, small α along with large β leads to low NIQE value. From the 3D curve in Fig. 12(b), we can find that the LOE value is the direct ratio to β . For the same β , changes of α also cause few impacts. On the contrary, in Fig. 12(c), IE value seems to be the direct ratio to α and does not change for the same α .

C. Gamma Correction Parameters – γ_1, γ_2

Fig. 13 demonstrates the impacts of gamma correction parameters, γ_1 and γ_2 , on the NIQE value, the LOE value and the IE value. As can be seen, lower γ_2 corresponds to

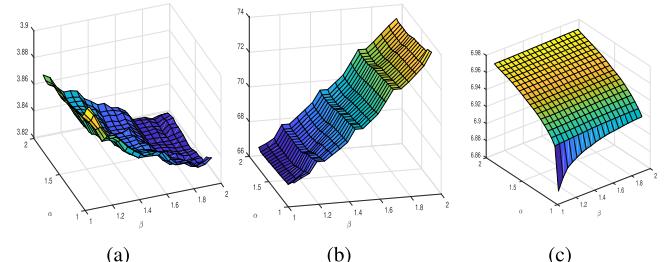


Fig. 12. Impacts of fractional-order parameters on average NIOE value and LOE value of 100 severely low-light images. (a) average NIQE value. (b) average LOE value. (c) average IE value.

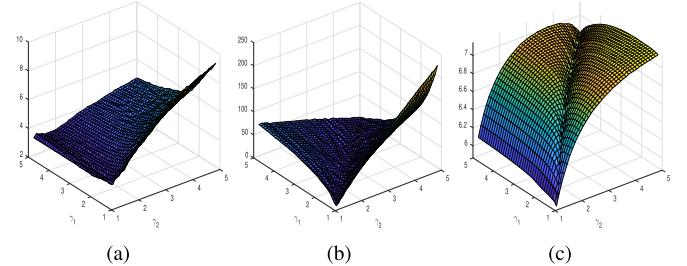


Fig. 13. Impacts of gamma correction parameters on average NIQE value and LOE value of 100 severely low-light images. (a) average NIQE value. (b) average LOE value. (c) average IE value.

TABLE II
AVERAGE VALUES ON 66 OTHER KINDS OF LOW-LIGHT IMAGES

| | [18] | [5] | [20] | [12] | [26] | Ours |
|------|-------------|-------------|--------|--------|-------|--------------|
| NIQE | 3.65 | 3.64 | 3.39 | 3.43 | 4.06 | 3.29 |
| LOE | – | 907.05 | 185.48 | 480.93 | 94.60 | 67.13 |
| IE | 7.38 | 7.38 | 7.36 | 6.98 | 7.04 | 7.36 |

lower NIQE and IE values while both lower or higher γ_1 and γ_2 contribute to good image quality.

VII. EXTENDED EXPERIMENTS

In this section, we first test the performance of the proposed method on other kinds of 66 low-light images, which are chosen from the Berkeley segmentation dataset [15] and the non-uniform illumination image dataset [23]. Then we exhibit results from other kinds of images including non-uniform illumination images, one side illumination images and backlit images. As can be seen from Fig. 14 (a), low-light image set A contains severely low-light image B. Naturally, properties which set B holds may not be suitable for set A, i.e., changing the gamma parameter pair (γ_1, γ_2) may contribute to enhance other kinds of low-light images. Here, we set γ_1 and γ_2 as 2.8 and 2.3 respectively.

Fig. 14 illustrates some results for comparison. Subjectively, our results have high contrast while there exists some over-enhancement which maybe resulted from the gamma correction on the reflectance. Actually, enhancement without correcting the reflectance may lead to low contrast results and worse evaluation indexes. On the other hand, our method has lower NIQE and LOE values, which can be observed from TABLE II. Therefore, in this extended experiment, our method has a trade-off between subjective and objective assessments.

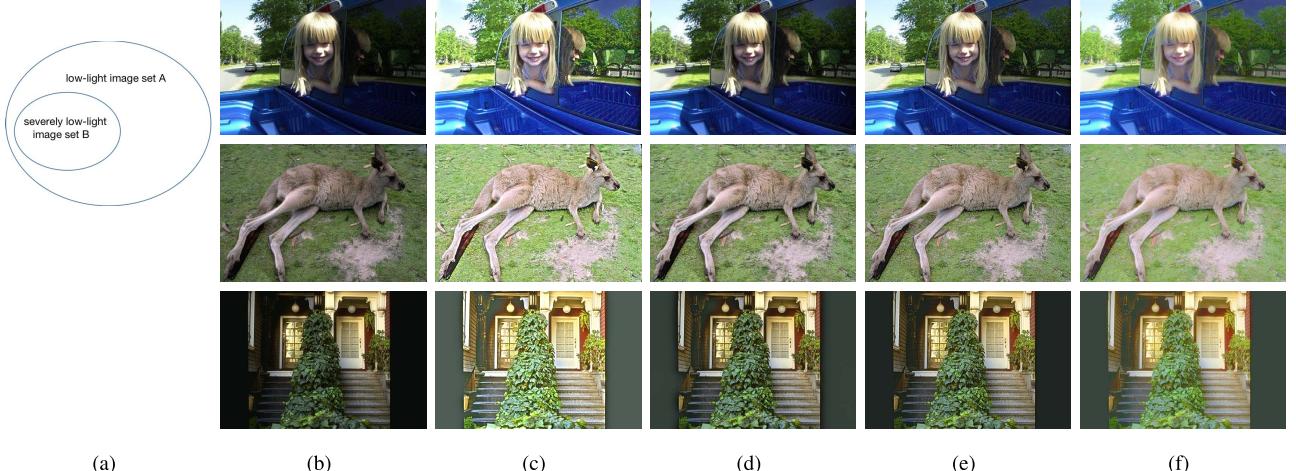


Fig. 14. Extended experiments: (a) relationship between low-light images and severely low-light images. (b) source images. (c) results from TVMR. (d) WVM. (e) LIME. (f) ours.



Fig. 15. Results from various light conditions. The first and third rows are the original images. The second and forth rows are the enhanced results.

Fig. 15 demonstrates some results from various light conditions. One can easily find that our method can effectively enhance the images while maintaining naturalness.

VIII. LIMITATIONS AND DISCUSSIONS

In this paper, we propose a novel Retinex-based fractional-order variational model for images with severely low light. We perform the decomposition directly in the image domain and the fractional-order gradient TV regularization terms on both the reflectance component and the illumination component to control the regularization extent more flexibly. Compared with other closely related Retinex-based methods, experimental results demonstrate the effectiveness of the proposed method. However, there are also some drawbacks. Firstly, each severely low-light image has its own characteristics, and the uniform fractional-order may not suitable for this task. Secondly, severely low-light images may contain complicated noise, and the noise may have impacts on the estimated results. Moreover, the fractional-order gradient may not model the real reflectance and the piecewise smoothness of the illumination.

In our future work, we will explore the adaptively fractional-order models. Besides, we will research the properties of the noise in dark areas and design reasonable noise suppressing methods to deal with the noise during the decomposition process. Moreover, studying more powerful mathematic tools

to model the real reflectance and the illumination is also necessary.

APPENDIX: PROOF OF THEOREM 2

Proof: Denote

$$\begin{aligned} \Phi_1 &:= (\nabla^\alpha)^T \Lambda_{d_1^k} - \Lambda_u^k + \sigma_1(R^{k+1} - u^k) + \sigma_3(\nabla^\alpha)^T \\ &\quad \times (d_1^k - \nabla^\alpha u^k), \end{aligned} \quad (73)$$

$$\begin{aligned} \Phi_2 &:= (\nabla^\beta)^T \Lambda_{d_2^k} - \Lambda_v^k + \sigma_2(L^{k+1} - v^k) + \sigma_4(\nabla^\beta)^T \\ &\quad (d_2^k - \nabla^\beta v^k). \end{aligned} \quad (74)$$

We can rewrite the updating formulas Eq. (51) - Eq. (54), Eq. (60) - Eq. (61) and Eq. (63) - Eq. (66) as the following equivalent forms

$$R^{k+1} - R^k = \mathcal{P}_1 \left(\frac{L^k I + \Lambda_u^k + \sigma_1 u^k}{(L^k)^2 + \sigma_1} \right) - R^k, \quad (75a)$$

$$L^{k+1} - L^k = \mathcal{P}_2 \left(\frac{R^{k+1} I + \Lambda_v^k + \sigma_2 v^k}{(R^{k+1})^2 + \sigma_2} \right) - L^k, \quad (75b)$$

$$(\sigma_1 + \sigma_3(\nabla^\alpha)^T \nabla^\alpha)(u^{k+1} - u^k) = \Phi_1, \quad (75c)$$

$$(\sigma_2 + \sigma_4(\nabla^\beta)^T \nabla^\beta)(v^{k+1} - v^k) = \Phi_2, \quad (75d)$$

$$d_1^{k+1} - d_1^k = \mathcal{S}(\nabla^\alpha u^{k+1} - \frac{\Lambda_{d_1^k}}{\sigma_3}, \frac{\lambda_1}{\sigma_3}) - d_1^k, \quad (75e)$$

$$\begin{aligned} d_2^{k+1} - d_2^k &= \mathcal{S}(\nabla^\beta v^{k+1} - \frac{\Lambda_{d_2^k}}{\sigma_4}, \frac{\lambda_2}{\sigma_4}) - d_2^k, \\ \Lambda_u^{k+1} - \Lambda_u^k &= \sigma_1(u^{k+1} - R^{k+1}), \\ \Lambda_v^{k+1} - \Lambda_v^k &= \sigma_2(v^{k+1} - L^{k+1}), \\ \Lambda_{d_1}^{k+1} - \Lambda_{d_1}^k &= \sigma_3(d_1^{k+1} - \nabla^\alpha u^{k+1}), \\ \Lambda_{d_2}^{k+1} - \Lambda_{d_2}^k &= \sigma_4(d_2^{k+1} - \nabla^\beta v^{k+1}). \end{aligned} \quad (75f) \quad (75g) \quad (75h) \quad (75i) \quad (75j)$$

Let $k \rightarrow \infty$ and note that

$$X^{k+1} - X^k \rightarrow 0 \quad Y^{k+1} - Y^k \rightarrow 0,$$

both sides of the above equations go to zero, that is,

$$\begin{aligned} 0 &\leftarrow \mathcal{P}_1 \left(\frac{L^k I + \Lambda_u^k + \sigma_1 u^k}{(L^k)^2 + \sigma_1} \right) - R^k, \\ 0 &\leftarrow \mathcal{P}_2 \left(\frac{R^{k+1} I + \Lambda_v^k + \sigma_2 v^k}{(R^{k+1})^2 + \sigma_2} \right) - L^k, \\ 0 &\leftarrow \Phi_1, \\ 0 &\leftarrow \Phi_2, \\ 0 &\leftarrow \mathcal{S}(\nabla^\alpha u^{k+1} - \frac{\Lambda_{d_1^k}}{\sigma_3}, \frac{\lambda_1}{\sigma_3}) - d_1^k, \\ 0 &\leftarrow \mathcal{S}(\nabla^\beta v^{k+1} - \frac{\Lambda_{d_2^k}}{\sigma_4}, \frac{\lambda_2}{\sigma_4}) - d_2^k, \\ 0 &\leftarrow u^k - R^k, \\ 0 &\leftarrow v^k - L^k, \\ 0 &\leftarrow d_1^k - \nabla^\alpha u^k, \\ 0 &\leftarrow d_2^k - \nabla^\beta v^k. \end{aligned} \quad (76a) \quad (76b) \quad (76c) \quad (76d) \quad (76e) \quad (76f) \quad (76g) \quad (76h) \quad (76i) \quad (76j)$$

Note that (76a), (76b), (76e) and (76f) are respectively equivalent to

$$0 \leftarrow L^k(R^k L^k - I) - \Lambda_u^k + \sigma_1(R^k - u^k) + \partial N_1(R^k), \quad (77a)$$

$$0 \leftarrow R^{k+1}(R^{k+1} L^k - I) - \Lambda_v^k + \sigma_2(L^k - v^k) + \partial N_2(L^k), \quad (77b)$$

$$0 \leftarrow \lambda_1 \partial \|d_1^k\|_1 + \Lambda_d^k + \sigma_3(d_1^k - \nabla^\alpha u^{k+1}), \quad (77c)$$

$$0 \leftarrow \lambda_2 \partial \|d_2^k\|_1 + \Lambda_d^k + \sigma_4(d_2^k - \nabla^\beta v^{k+1}). \quad (77d)$$

Using (76g) - (76j), the corresponding terms in (76) and (77) can be eliminated when $k \rightarrow \infty$. For any limit point \hat{X} of sequence $\{X^k\}$, there exists a subsequence $\{X^{k_i}\}$ converging to \hat{X} . Besides, the boundedness of $\{Y^k\}$ implies there existing a subsequence $\{X^{k_j}\}$ converging to some limit point \hat{Y} . So (\hat{X}, \hat{Y}) is a limit point of the sequence $\{X^k, Y^k\}$. Letting $k \rightarrow \infty$, it's easy to check it that $\{X^k, Y^k\}$ is a KKT point of the problem (46). Hence, (\hat{R}, \hat{L}) is a KKT point of the problem (46) and further the original problem (33). This completes the proof. \square

ACKNOWLEDGMENT

The authors would like to thank W. Wang, X. Fu and M. Li for sharing the codes of [18], [5] and [12] respectively. They also thank Martin *et al.*, Chen *et al.* and Wang *et al.* for offering the Berkeley Segmentation dataset [15], the Pku-EQAQ dataset [4] and the non-uniform illumination image dataset [23], respectively.

REFERENCES

- [1] R. H. Bartels and G. W. Stewart, "Solution of the matrix equation $AX + XB = C$," *Commun. ACM*, vol. 15, no. 9, pp. 820–826, Sep. 1972.
- [2] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Nov. 2010.
- [3] J. C. Caicedo, A. Kapoor, and S. B. Kang, "Collaborative personalization of image enhancement," in *Proc. CVPR*, Jun. 2011, pp. 249–256.
- [4] Z. Chen, T. Jiang, and Y. Tian, "Quality assessment for comparing image enhancement algorithms," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 3003–3010.
- [5] X. Fu, D. Zeng, Y. Huang, X.-P. Zhang, and X. Ding, "A weighted variational model for simultaneous reflectance and illumination estimation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 2782–2790.
- [6] Z. Gu, F. Li, and X.-G. Lv, "A detail preserving variational model for image Retinex," *Appl. Math. Model.*, vol. 68, pp. 643–661, Apr. 2019.
- [7] B. K. P. Horn, "Determining lightness from an image," *Comput. Graph. Image Process.*, vol. 3, no. 4, pp. 277–299, Dec. 1974.
- [8] R. Kimmel, M. Elad, D. Shaked, R. Keshet, and I. Sobel, "A variational framework for Retinex," *Int. J. Comput. Vis.*, vol. 52, no. 1, pp. 7–23, Apr. 2003.
- [9] E. H. Land, "The Retinex theory of color vision," *Sci. Amer.*, vol. 237, no. 6, pp. 108–129, Dec. 1977.
- [10] E. H. Land, "Recent advances in Retinex theory and some implications for cortical computations: Color vision and the natural image," *Proc. Nat. Acad. Sci. USA*, vol. 80, no. 16, pp. 5163–5169, Aug. 1983.
- [11] E. H. Land and J. J. McCann, "Lightness and Retinex theory," *J. Opt. Soc. Amer.*, vol. 61, no. 1, pp. 1–11, 1971.
- [12] M. Li, J. Liu, W. Yang, X. Sun, and Z. Guo, "Structure-revealing low-light image enhancement via robust Retinex model," *IEEE Trans. Image Process.*, vol. 27, no. 6, pp. 2828–2841, Jun. 2018.
- [13] J. Liang and X. Zhang, "Retinex by higher order total variation ℓ^1 decomposition," *J. Math. Imag. Vis.*, vol. 52, no. 3, pp. 345–355, Jul. 2015.
- [14] W. Ma and S. Osher, "A TV Bregman iterative model of Retinex theory," *Inverse Problems Imag.*, vol. 6, no. 4, pp. 697–708, Nov. 2012.
- [15] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *Proc. 8th IEEE Int. Conf. Comput. Vis. (ICCV)*, Vancouver, BC, Canada, Jul. 2001, pp. 416–423.
- [16] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a 'completely blind' image quality analyzer," *IEEE Signal Process. Lett.*, vol. 20, no. 3, pp. 209–212, Mar. 2013.
- [17] 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, Nov. 2010.
- [18] M. K. Ng and W. Wang, "A total variation model for Retinex," *SIAM J. Imag. Sci.*, vol. 4, no. 1, pp. 345–365, 2011.
- [19] K. Oldham and J. Spanier, *The Fractional Calculus Theory and Applications of Differentiation and Integration to Arbitrary Order*, vol. 111. Amsterdam, The Netherlands: Elsevier, 1974.
- [20] S. Park, S. Yu, B. Moon, S. Ko, and J. Paik, "Low-light image enhancement using variational optimization-based Retinex model," *IEEE Trans. Consum. Electron.*, vol. 63, no. 2, pp. 178–184, May 2017.
- [21] I. Podlubny, A. Chechkin, T. Skovranek, Y. Chen, and B. M. V. Jara, "Matrix approach to discrete fractional calculus II: Partial fractional differential equations," *J. Comput. Phys.*, vol. 228, no. 8, pp. 3137–3153, May 2009.
- [22] H. Wang and N. Du, "Fast solution methods for space-fractional diffusion equations," *J. Comput. Appl. Math.*, vol. 255, pp. 376–383, Jan. 2014.
- [23] S. Wang, J. Zheng, H.-M. Hu, and B. Li, "Naturalness preserved enhancement algorithm for non-uniform illumination images," *IEEE Trans. Image Process.*, vol. 22, no. 9, pp. 3538–3548, Sep. 2013.
- [24] W. Wang and C. He, "A variational model with barrier functionals for Retinex," *SIAM J. Imag. Sci.*, vol. 8, no. 3, pp. 1955–1980, 2015.
- [25] Y. Xu, W. Yin, Z. Wen, and Y. Zhang, "An alternating direction algorithm for matrix completion with nonnegative factors," *Frontiers Math. China*, vol. 7, no. 2, pp. 365–384, Apr. 2012.

- [26] Z. Ying, G. Li, Y. Ren, R. Wang, and W. Wang, "A new low-light image enhancement algorithm using camera response model," in *Proc. IEEE Int. Conf. Comput. Vis. Workshops (ICCVW)*, Oct. 2017, pp. 3015–3022.
- [27] J. Zhang and K. Chen, "A total fractional-order variation model for image restoration with nonhomogeneous boundary conditions and its numerical solution," *SIAM J. Imag. Sci.*, vol. 8, no. 4, pp. 2487–2518, 2015.



Zhihao Gu received the M.S. degree from East China Normal University, Shanghai, China, in 2019. He is currently pursuing the Ph.D. degree with the Department of Computer Science and Engineering, Shanghai Jiao Tong University, China. His research interests include image processing and object detection.



Fang Li received the Ph.D. degree in mathematics from East China Normal University, Shanghai, China, in 2007. She is currently a Professor with the School of Mathematical Sciences, East China Normal University. Her research interests include variational methods and deep learning methods in image processing and computer vision.



Faming Fang received the Ph.D. degree in computer science from East China Normal University, Shanghai, China, in 2013. He is currently an Associate Professor with the School of Computer Science and Technology, East China Normal University. His main research area is image processing using the variational methods, PDEs, and deep learning.



Guixu Zhang received the Ph.D. degree from the Institute of Modern Physics, Chinese Academy of Sciences, Lanzhou, China, in 1998. He is currently a Professor with the Department of Computer Science and Technology, East China Normal University, Shanghai, China. His research interests include hyperspectral remote sensing, image processing, and artificial intelligence.