DOI: 10.1111/cgf.14960
Pacific Graphics 2023
R. Chaine, Z. Deng, and M. H. Kim
(Guest Editors)

COMPUTER GRAPHICS *forum*
Volume 42 (2023), Number 7

Enhancing Low-Light Images: A Variation-based Retinex with Modified Bilateral Total Variation and Tensor Sparse Coding

Weipeng Yang[1], Hongxia Gao[†,1,2], Wenbin Zou[1], Shasha Huang[1], Hongsheng Chen[1], Jianliang Ma[1,3]

[11]School of Automation Science and Engineering, South China University of Technology, Guangzhou, China
[2]Research Center for Brain-Computer Interface, Pazhou Laboratory, Guangzhou, China
[3]KUKA Robotics Guangdong Co., Ltd., Foshan, China

Abstract

Low-light conditions often result in the presence of significant noise and artifacts in captured images, which can be further exacerbated during the image enhancement process, leading to a decrease in visual quality. This paper aims to present an effective low-light image enhancement model based on the variation Retinex model that successfully suppresses noise and artifacts while preserving image details. To achieve this, we propose a modified Bilateral Total Variation to better smooth out fine textures in the illuminance component while maintaining weak structures. Additionally, tensor sparse coding is employed as a regularization term to remove noise and artifacts from the reflectance component. Experimental results on extensive and challenging datasets demonstrate the effectiveness of the proposed method, exhibiting superior or comparable performance compared to state-of-the-art approaches. Code, dataset and experimental results are available at https://github.com/YangWeipengscut/BTRetinex.

CCS Concepts

• ***Computing methodologies*** → *Image processing; Low-level-vision tasks;*

1. Introduction

The Retinex theory is a color vision model that aims to explain how humans perceive color under varying conditions of illuminance [Lan77]. Specifically, the Retinex theory aims at decomposing an observed image $\boldsymbol{S}$ into illuminance component $\boldsymbol{L}$ and reflectance component $\boldsymbol{R}$ via $\boldsymbol{S} = \boldsymbol{L} \circ \boldsymbol{R}$, where $\circ$ represents element-wise multiplication. The illuminance component $\boldsymbol{L}$ represents the influence of the light source intensity and color on the object color, while the reflectance component $\boldsymbol{R}$ represents the inherent color information of the object itself [CXG*17].The Retinex theory is used in many image processing tasks, such as low-light image enhancement [RJW96, WZHL13, CXG*17, LL22], and color correction [FZH*16b, XHR*20].

With the increasing popularity of the Retinex model, numerous researchers have discovered the valuable properties of the illuminance and reflectance components, which can effectively address the challenges of low-light image enhancement and color correction. These models can be broadly categorized into three types: classical methods, variation-based methods and learning-based methods. Early classical Retinex methods [BW86, Lan83] were based on the assumption that the reflectance component could be computed using random paths. However, these methods were computationally expensive and lacked stability in practical applications. To overcome these limitations, researchers [MPS10] utilized the solution of the Retinex model's differential equation, satisfying the discrete Poisson equation, to efficiently estimate the reflectance using fast Fourier transform. However, due to the piecewise smooth gradients of the illuminance component, its structure could suffer from degradation. Therefore, corresponding single-scale Retinex [JRW97b] and multi-scale Retinex [RJW96,JRW97a] methods were proposed. These methods typically assume that the illuminance component is piecewise smooth, while the reflectance component is non-smooth, resulting in improved model performance. However, due to the lack of appropriate edge constraints, these methods often exhibit halo artifacts and pseudo-shadows around edges.

To improve the application of the smoothness assumption in estimating the illuminance component within the Retinex model, researchers [MMOC11, PDCRM05] have introduced variation-based methods for decomposing the illuminance and reflectance based on the Retinex model. Estimating illuminance and reflectance from a single observed image is an ill-posed problem. The variation-based Retinex method offers a promising solution by estimating piecewise smooth illuminance and piece-wise continuous reflectance, then adjusts the illuminance and multiplies it with the reflectance to obtain the enhanced result. In the context of the variation-based Retinex method, the primary design methods for the regularization

† Corresponding author

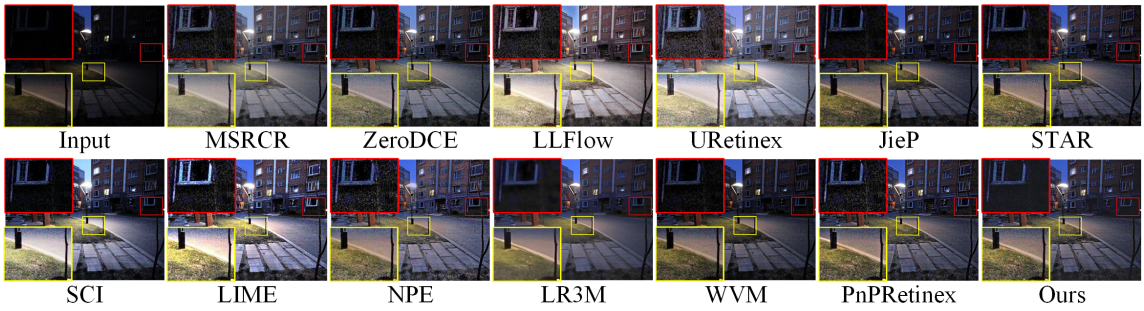

Figure 1: *Comparative analysis of various representative low-light image enhancement methods, including MSRCR [RJW96], ZeroDCE [GLG*20], LLFlow [WWY*22], URetinex [WWZ*22], JiePP [CXG*17], STAR [XHR*20], SCI [MML*22], LIME [GLL16], NPE [WZHL13], LR3M [RYCL20], WVM [FZH*16b] and PnPRetinex [LL22].*

terms of the illuminance component and reflectance component are as follows:

1) In order to obtain piece-wise smooth illuminance components, researchers have proposed different regularization terms to distinguish the details and structures of the illuminance components. For instance, WVM [FZH*16b] method utilizes the L_2 norm of the illuminance component gradients, while the SRRR [LLY*18] and LR3M [RYCL20] methods employ the total variation of the illuminance component gradients. The Jiep [CXG*17] method utilizes the local variation deviation of the illuminance component, and the STAR [XHR*20] method utilizes the exponentiated mean local variance of the illuminance component. The PnPRetinex [LL22] method combines the fractional norm and guided filter [HST12] to constrain the illuminance component. However, these methods have certain limitations. On one hand, SRRR, LR3M, Jiep and STAR methods produce illuminance components that contain excessive details, resulting in reduced contrast in the enhanced images. On the other hand, WVM and PnPRetinex methods excessively smooth the intense or structures of the illuminance components, leading to inaccurate brightness estimation in the enhanced images and subsequent loss of color information.
2) Furthermore, devising a regularization term to eliminate noise and artifacts in the reflectance component while preserving its details remains a challenging task. Methods [FZH*16b, FLZ*15, LLY*18, CXG*17, XHR*20, LL22, RYCL20] utilize only *local* priors, such as L_1, L_2 or L_p norm ($0<p<1$), to penalize the reflectance gradient. In order to better distinguish noise and details, methods [FZH*16b, XHR*20, LL22] also introduce weighted strategies based on local variance or gradient into the norm of the reflectance gradient. It's worth noting that these methods are effective only when noise levels are low. In the case of intense noise, these methods based solely on local prior of the reflectance gradient either have difficulty reducing noise or result in over-smoothing of details and artifacts. Later on, Ren *et al.* [RYCL20] were the first to use the low-rank matrix prior based on *non-local* self-similarity to reduce intense noise and artifacts in the reflectance. However, they uniformly shrunk all singular values of the reflectance, which ignored the physical meaning of singular values with varying amplitudes. Consequently, although this method can reduced noise somewhat, it also led to excessive smoothing of image details. In addition, all of these methods modeled the prior of color image in vector space, which obviously overlooks the multidimensional structure of color images and inevitably results in structural losses and distortions [CYZ*20].

With the remarkable achievements of deep learning across various domains, [WWYL18] pioneered the application of deep learning to the Retinex model by proposing a two-stage Retinex-based method known as Retinex-Net. Retinex-Net comprises a decomposition module and an illuminance adjustment module. Drawing inspiration from Retinex-Net, Zhang et al. introduced two improved methods called KinD [ZZG19] and KinD++ [ZGM*21]. Expanding on these advancements, [WCZ*19] proposed a progressive Retinex network consisting of an illuminance module network for estimating illuminance and an noise module network for estimating noise level. These two sub-networks operate progressively until stable results are achieved. More recently, Wu et al. [WWZ*22] introduced a novel deep unfolding network based on Retinex known as URetinex, aiming to further integrate the benefits of variation-based and learning-based approaches. Nevertheless, it is worth noting that deep learning-based methods typically necessitate a significant amount of paired training data. Moreover, when the statistical characteristics of the test image diverge from those of the training images, the enhanced results often exhibit visual artifacts, color distortion or excessive enhancement.

To address the limitations of existing variation-based Retinex methods in low-light image enhancement tasks, this paper presents the following main contributions:

1) This paper presents a modified bilateral total variation method that effectively removes textures while preserves weak structures in the illuminance component of low-light images.
2) To address the issue of noise and artifacts in the reflectance component, the paper proposes the utilization of tensor sparse coding as a regularization term. This technique successfully suppresses noise and artifacts, resulting in a cleaner and more visually appealing reflectance component, while effectively preserving the details of the image.
3) The effectiveness of the proposed method is demonstrated through comprehensive evaluation on extensive and challenging

datasets. It outperforms or performs comparably to state-of-the-art approaches in enhancing low-light images.

2. BACKGROUND

2.1. Notations

A tensor can be treated as a multi-indexed array, with its order determined by the number of dimensions or modes. In this paper, a tensor of N-order is denoted by Euler script letter, e.g., $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$. A matrix is a tensor of 2-order, denoted by boldface captital letter, e.g., $\boldsymbol{X}$, and a vector is a tensor of 1-order, denoted as $\boldsymbol{x}$.

2.2. Variation-based Retinex Model

Traditional variation-based Retinex models are commonly formulated using matrix representations. For example, they may separately process the red, green and blue channels of a color image [RJW96], or transform the color image into the HSV color space and operate only on the luminance channel [RYCL20, XHR*20, LL22, CXG*17]. The matrix representation of the Retinex model can be expressed as follows:

$$\boldsymbol{S} = \boldsymbol{L} \circ \boldsymbol{R}, \tag{1}$$

where $\circ$ represents element-wise multiplication. The illuminance component $\boldsymbol{L}$ describes the distribution and intensity of light in a scene, which is related to the lighting conditions in the environment. It typically consists of low-frequency components that represent the overall illuminance variations in the image. In contrast, the reflectance component $\boldsymbol{R}$ reflects the surface characteristics of objects in terms of light reflectance and is associated with material and texture properties. It usually contains high-frequency components that capture the local details in the image.

Estimating illuminance and reflectance from a single observed image is an ill-posed problem. The variation Retinex method offers a promising solution by estimating piece-wise smooth illuminance and piece-wise continuous reflectance. This is achieved by introducing specific regularization terms, such as total variation [CXG*17] and low-rank priors [RYCL20], to enhance the accuracy and robustness of the estimation. The general framework of the variation-based Retinex model is outlined as follows:

$$\min_{\boldsymbol{L},\boldsymbol{R}} \|\boldsymbol{S} - \boldsymbol{L} \circ \boldsymbol{R}\|_F^2 + \alpha\Phi_1(\boldsymbol{L}) + \beta\Phi_2(\boldsymbol{R}), \tag{2}$$

where $\|\boldsymbol{S} - \boldsymbol{L} \cdot \boldsymbol{R}\|_F^2$ represents the fidelity term, while $\Phi_1(\boldsymbol{L})$ and $\Phi_2(\boldsymbol{R})$ denote the constraint terms imposed on $\boldsymbol{l}$ and $\boldsymbol{R}$, respectively. The parameters α and β serve as balance parameters, which are utilized to adjust the weighting of these constraint terms during the estimation process. In Figure 2, we illustrate the decomposition results of the latest variation-based Retinex methods.

2.3. Modified Bilateral Total Variation and Tensor Sparse Coding

Low-light images often suffer from issues like noise and artifacts, which can interfere with the decomposition process and pose challenges in accurately extracting the illuminance and reflectance components. To overcome this challenge, we begin by redesigning the bilateral total variation to constrain the illuminance component, smoothing out fine textures while preserving structural information. Additionally, we employ tensor sparse priors to constrain the reflectance component, aiming to eliminate noise and artifacts within it. Therefore, we jointly employ the bilateral total variation prior and tensor sparse prior to simultaneously estimating the illuminance $\mathcal{L} \in \mathbb{R}^{m \times m \times 3}$ and reflectance $\mathcal{R} \in \mathbb{R}^{m \times m \times 3}$ from the observed low-light image $\mathcal{S} \in \mathbb{R}^{m \times m \times 3}$. Here, m, m and 3 respectively represent the height, width and number of channels of the image.

2.3.1. Modified Bilateral Total Variation

The primary objective of introducing the Bilateral Total Variation (BTV) [HXXC22] was to facilitate texture smoothing and preserve image structure. In contrast to conventional methods like Total Variation (TV) [ROF92], Relative Total Variation (RTV) [XYXJ12] or Bilateral Texture Filtering (BTF) [CLKL14a], BTV exhibits superior discrimination between image texture and structure. It effectively smooths image textures while simultaneously preserving weak structures. The presence of these weak structures can be attributed to variations in illuminance or changes in surface characteristics of the scene, necessitating their retention within the illuminance component.

In this section, we propose a modified Bilateral Total Variation (mBTV), designed to enhance the robustness of texture and structure discrimination in complex scenarios, such as the coexistence of textures and weak structures within a region or the presence of noise. It is denoted as

$$mBTV(\mathcal{L})_i = \left\| \frac{\mathscr{D}_h(i)}{GB_h(i) + \varepsilon} \right\|_1 + \left\| \frac{\mathscr{D}_v(i)}{GB_v(i) + \varepsilon} \right\|_1, \tag{3}$$

where i represents the spatial position coordinates, ε is a small constant introduced to prevent division by zero.

The *windowed total variation*, denoted as $\mathscr{D}_{h/v}(i)$, is defined as the weighted sum of the absolute of spatial differences within a local window. The expression for its horizontal direction x or vertical direction y is given as follows:

$$\mathscr{D}_{h/v}(i) = \sum_{j \in \Omega(i)} g_{k_s}(i,j) \cdot |(\partial_{h/v}\mathcal{L})_j|, \tag{4}$$

where $\Omega(i)$ represents a local window at pixel $i = (x_i, y_i)$ with the size of $(2 * k_s + 1) \times (2 * k_s + 1)$, k_s means the spatial scale. The weight $g_{k_s}(i,j)$ at pixel $j = (x_j, y_j)$ is determined by a Gaussian function with a standard deviation of k_s, written as

$$g_{k_s}(i,j) = \frac{1}{\sqrt{2\pi}k_s} exp\left(-\frac{(x_i - x_j)^2 + (y_i - y_j)^2}{2k_s^2} \right). \tag{5}$$

$\mathscr{D}_{h/v}(i)$ can only differentiate between smooth regions and oscillatory areas. If $\mathscr{D}_{h/v}(i)$ is solely used as a constraint term for the illuminance component $\mathcal{L}$, it would be challenging to distinguish textures and weak structures . As a result, both textures and weak structures would be subject to smoothing, leading to inaccurate estimation of the illuminance component, similar to the methods [FLZ*15, CXG*17].

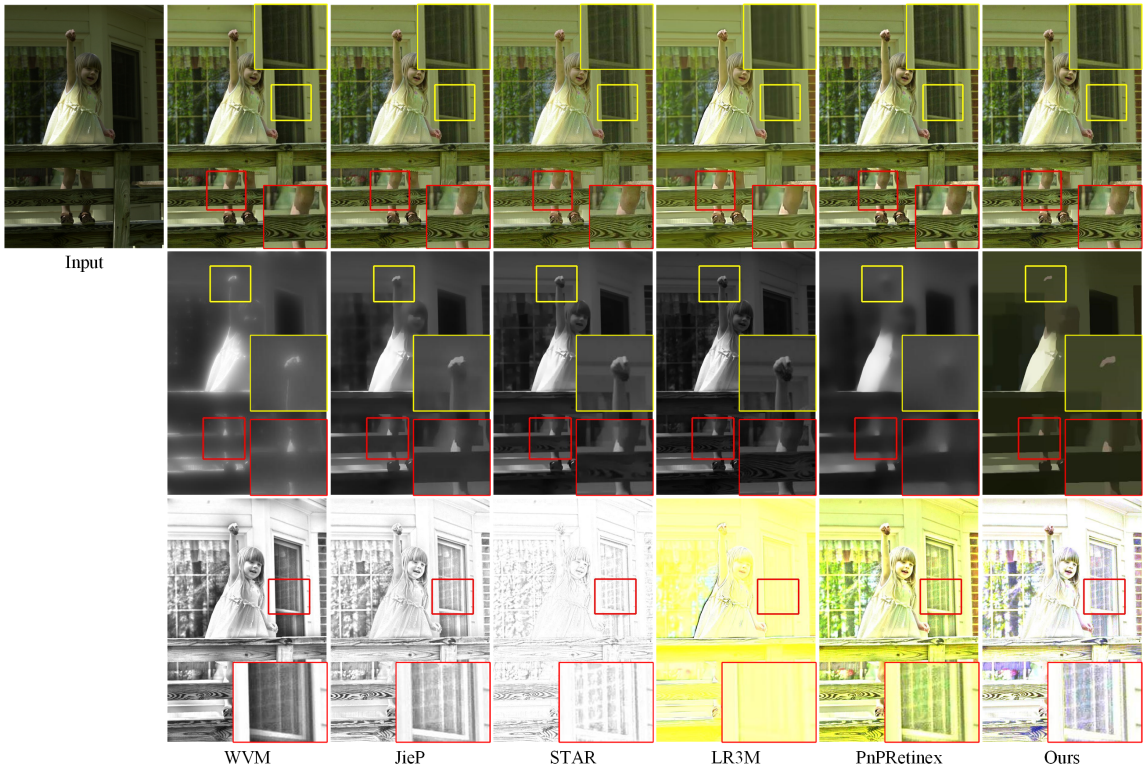

Figure 2: *Comparative analysis of the enhancement and decomposition results between our proposed model and the latest variation-based Retinex methods, namely WVM [FZH*16b], JieP [CXG*17], STAR [XHR*20], LR3M [RYCL20] and PnPRetinex [LL22], under low noise level. Row 1 represents the enhanced results, while rows 2 and 3 correspond the illuminance and reflectance components, respectively.*

To overcome this limitation, [HXXC22] proposed a *guidance bilateral variation* $GB_{h/v}(i)$ based on bilateral filtering, which can be expressed as follows:

$$GB_{h/v}(i) = \left\| (\partial_{h/v}(f_{GB}(\mathcal{L},\mathcal{G})))_i \right\|_1, \tag{6}$$

$$f_{GB}(\mathcal{L},\mathcal{G})_i = \frac{1}{k_n} \sum_{j\in\Omega(i)} f_{k_s}(i,j)\cdot f_{k_r}(\mathcal{G}_i - \mathcal{G}_j)\cdot \mathcal{L}_i, \tag{7}$$

where $\partial_{h/v}$ is the partial derivatives in horizontal or vertical direction, f_{GB} is the guidance bilateral filter, k_n serves as a normalization factor. The values of $f_{k_s}(i,j)$ and $f_{k_r}(\mathcal{G}_i - \mathcal{G}_j)$ are calculated based on the Gaussian functions of the spatial and range distance between two pixels, namely $i = (x_i, y_i)$ and $j = (x_j, y_j)$, respectively. They are defined as follows:

$$f_{k_s}(i,j) = exp\left(-\frac{(x_i - x_j)^2 + (y_i - y_j)^2}{2k_s^2}\right), \tag{8}$$

$$f_{k_r}(\mathcal{G}_i - \mathcal{G}_j) = exp\left(-\frac{(\mathcal{G}_i - \mathcal{G}_j)^2 + (\mathcal{G}_i - \mathcal{G}_j)^2}{2k_r^2}\right). \tag{9}$$

The range parameter k_r controls the sensitivity of f_{GB} to the structure. $\mathcal{G}$ represents the guidance map, obtained by applying Gaussian filtering to the smoothed image in [HXXC22]. To handle complex scenarios, such as the coexistence of textures and weak structures within a region or the presence of noise, this paper directly calculates the guidance map based on the *modified relative total variation* (mRTV), as elaborated in [CLKL14b].

To enhance the discriminative power between textures and edges, we combine terms $\mathscr{D}_{h/v}(i)$ and $GB_{h/v}(i)$ to form our novel regularization term, namely the modified bilateral total variation, as shown in Eq. (3).

2.3.2. Tensor Sparse Coding

Sparse dictionary learning [OF97, AEB06, EA06] has emerged as a powerful technique for image denoising. It usually involves learning an over-complete dictionary of atoms from a set of training images and utilizing it to efficiently represent and reconstruct clean images from noisy observations. The underlying principle of dictionary learning is based on the assumption that images can be sparsely represented in terms of a few atoms from the over-complete dictionary.

In sparse dictionary learning, the goal is to find a dictionary matrix $\boldsymbol{D}$ and a sparse coding vector $\boldsymbol{x}$ that best represents the clean image from noisy observations $\boldsymbol{z}$. The formulation can be expressed as follows:

$$\min_{\boldsymbol{D},\boldsymbol{x}} \|\boldsymbol{z}-\boldsymbol{D}\boldsymbol{x}\|_F^2 + \lambda \|\boldsymbol{x}\|_0, \tag{10}$$

where $||\cdot||_F$ denotes the the Frobenius norm [CYZ*20], $||\cdot||_0$ represents the sparsity-promoting L_0 norm, and λ is a regularization parameter that controls the sparsity level. The term $||\boldsymbol{z}-\boldsymbol{D}\boldsymbol{x}||_F^2$ measures the fidelity of the reconstruction by minimizing the squared difference between the noisy image $\boldsymbol{z}$ and the reconstructed image $\boldsymbol{D}\boldsymbol{x}$.

However, matrix-based dictionary learning assumes that data can be adequately represented using a matrix, which may not be suitable for high-dimensional data such as images, videos or hyperspectral data. When applied directly to high-dimensional data, matrix dictionary learning fails to exploit the inherent multiway structure and correlations among different dimensions, resulting in suboptimal representations.

Tensor dictionary learning [DWL*12, QHJ15, DHQ*18] extends the concept of dictionary learning from matrices to higher-order tensors, allowing for the modeling of more complex and structured data. This is particularly beneficial when dealing with multidimensional data such as color images, videos, and hyperspectral images, where the relationship between different modes of data can be effectively captured using tensor representations.

3. Our method

In this section, we propose a new variation-based Retinex model based on modified Bilateral total variation and Tensor sparse coding, referred to as BTRetinex, for simultaneously estimating the illuminance $\mathcal{L}$ and reflectance $\mathcal{R}$ from the observed image $\mathcal{S}$ based on $\mathcal{S}=\mathcal{L}\circ\mathcal{R}$.

The proposed BTRetinex model is represented as follows:

$$\begin{aligned}\arg\min_{\mathcal{L},\mathcal{R},\boldsymbol{D},\boldsymbol{C}} &\|\mathcal{S}-\mathcal{L}\circ\mathcal{R}\|_F^2 + \lambda_1 \sum_{i=1}^{3mn} mBTV(\mathcal{L})_i + \frac{\lambda_2}{2}\|\mathcal{R}-\mathcal{P}(\boldsymbol{D}\boldsymbol{C})\|_F^2 \\ &+\lambda_3\|\boldsymbol{C}\|_0 \\ &s.t.\boldsymbol{D}^T\boldsymbol{D}=\boldsymbol{I},\end{aligned} \tag{11}$$

wherein, the first term represents the data fidelity term. The second term corresponds to the regularization term of the illuminance component, which is based on modified bilateral total variation. Its objective is to achieve adequate piece-wise smoothness while preserving weak structures in the illuminance component. The third and fourth terms represent the regularization terms of the reflectance component, employing tensor sparse coding. The primary objective is to eliminate noise and artifacts in the reflectance component while preserving important details. λ_1, λ_2, λ_3 are regularization parameters to balance the terms. $\mathcal{P}$ is an operation used to aggregate matrix into a tensor [DHQ*18]. $\boldsymbol{I}$ is the identity matrix. In order to enhance the efficiency of the proposed model, this paper adopts an orthogonal dictionary $\boldsymbol{D}$, similar to the approach employed in [BCJ13, CJSY14].

3.1. Solution for Our Model

The objective function in Eq. (11) is separable with respect to variables $\mathcal{L}$, $\mathcal{R}$, $\boldsymbol{D}$ and $\boldsymbol{C}$, allowing us to employ an alternating optimization approach for solving it. Specifically, we optimize one variable while keeping the other three fixed. We initialize the illuminance component $\mathcal{L}^0$ as $\mathcal{S}$ and the reflectance component $\mathcal{R}^0$ as $\mathcal{I}$, where $\mathcal{I}$ represents a unit tensor.

All the sub-problems in the $(k+1)$-th iteration are listed in the following equation:

$$\begin{cases}\mathcal{L}^{k+1} = \arg\min_{\mathcal{L}} \left\|\mathcal{S}-\mathcal{L}\circ\mathcal{R}^k\right\|_F^2 + \lambda_1 \sum_{i=1}^{3mn} mBTV(\mathcal{L})_i \\ \mathcal{R}^{k+1} = \arg\min_{\mathcal{R}} \left\|\mathcal{S}-\mathcal{L}^{k+1}\circ\mathcal{R}\right\|_F^2 + \frac{\lambda_2}{2}\left\|\mathcal{R}-\mathcal{P}^{-1}(\boldsymbol{D}^k\boldsymbol{C}^k)\right\|_F^2 \\ \boldsymbol{D}^{k+1} = \arg\min_{\boldsymbol{D}^T\boldsymbol{D}=\boldsymbol{I}} \left\|\mathcal{R}^{k+1}-\mathcal{P}^{-1}(\boldsymbol{D}^k\boldsymbol{C})\right\|_F^2 \\ \boldsymbol{C}^{k+1} = \arg\min_{\boldsymbol{C}} \frac{\lambda_2}{2}\left\|\mathcal{R}^{k+1}-\mathcal{P}^{-1}(\boldsymbol{D}^{k+1}\boldsymbol{C})\right\|_F^2 + \lambda_3\|\boldsymbol{C}\|_0.\end{cases} \tag{12}$$

For the sake of conciseness, the superscript k or $k+1$ denoting the iteration number is partially omitted in the subsequent solving process.

(1) ***Update*** $\mathcal{L}$: The corresponding optimization sub-problem for $\mathcal{L}^{k+1}$ is

$$\begin{aligned}\mathcal{L}^{k+1} &= \arg\min_{\mathcal{L}} \|\mathcal{S}-\mathcal{L}\circ\mathcal{R}\|_F^2 + \lambda_1 \sum_{i=1}^{3mn} mBTV(\mathcal{L})_i \\ &= \arg\min_{\mathcal{L}} \|\boldsymbol{s}-\boldsymbol{l}\circ\boldsymbol{r}\|_F^2 \\ &+\lambda_1\left(\boldsymbol{l}^T\boldsymbol{M}_h^T\boldsymbol{W}_h\boldsymbol{M}_h\boldsymbol{l}+\boldsymbol{l}^T\boldsymbol{M}_v^T\boldsymbol{W}_v\boldsymbol{M}_v\boldsymbol{l}\right),\end{aligned} \tag{13}$$

where $\boldsymbol{M}_h$, $\boldsymbol{M}_v$ are Toeplitz matrices from discrete gradient operators with forward difference. The vectors $\boldsymbol{s}$, $\boldsymbol{l}$ and $\boldsymbol{r}$ represent the vector representations of $\mathcal{S}$, $\mathcal{L}$ and $\mathcal{R}$, respectively. The weight vectors w_h, w_v are computed according to the methods described in [XYXJ12] or [HXXC22], respectively. The expressions for the them at position i are as follows:

$$\boldsymbol{w}_{h/v}^i = \left(\sum_{j\in\Omega(i)} \frac{g_{k_s}(i,j)}{GB_{h/v}(i)+\varepsilon}\right)\frac{1}{|(\partial_{h/v}\mathcal{L})_i|+\varepsilon}, \quad i=1,2,\cdots,(mn) \tag{14}$$

It should be noted that for the color illuminance component, we need to calculate the weight vectors $\boldsymbol{w}_h$ and $\boldsymbol{w}_v$ separately for each channel c ($c=1,2,3$). To simplify the equation representation, we have omitted the channel indicators here and directly represented $\boldsymbol{w}_h$ and $\boldsymbol{w}_v$ as the concatenation of the weight vectors for the three channels. The matrices $\boldsymbol{W}_h$ and $\boldsymbol{W}_h$ represent the diagonalization of matrices $\boldsymbol{w}_h$ and $\boldsymbol{w}_v$, respectively.

By taking the derivative of Eq. (13) with respect to $\boldsymbol{l}$ and setting the derivative equal to 0, we obtain

$$((\boldsymbol{D}_r\circ\boldsymbol{D}_r)+\lambda_1\boldsymbol{Q})\,\boldsymbol{l}=\boldsymbol{D}_r\boldsymbol{s}, \tag{15}$$

where $\boldsymbol{D}_r$ is the diagonalization matrix of $\boldsymbol{r}$, $\boldsymbol{Q} = \boldsymbol{M}_h^T \boldsymbol{W}_h \boldsymbol{M}_h + \boldsymbol{M}_v^T \boldsymbol{W}_v \boldsymbol{M}_v$.

Then, the solution for $\boldsymbol{l}$ can be transformed into solving a series of linear equation systems:

$$\boldsymbol{l} = ((\boldsymbol{D}_r \circ \boldsymbol{D}_r) + \lambda_1 \boldsymbol{Q})^{-1} (\boldsymbol{D}_r \boldsymbol{s}). \tag{16}$$

After obtaining the solution for $\boldsymbol{l} \in \mathbb{R}^{3mn \times 1}$, we reshape it into $\mathcal{L} \in \mathbb{R}^{m \times n \times 3}$.

(2) ***Update*** $\mathcal{R}$: The corresponding optimization sub-problem for $\mathcal{R}^{k+1}$ is

$$\mathcal{R}^{k+1} = \arg\min_{\mathcal{R}} \|\mathcal{S} - \mathcal{L} \circ \mathcal{R}\|_F^2 + \frac{\lambda_2}{2} \|\mathcal{R} - \mathcal{P}(\boldsymbol{DC})\|_F^2. \tag{17}$$

This is a least squares problem and its closed-form solution is given by

$$(2(\boldsymbol{D}_l \circ \boldsymbol{D}_l) + \boldsymbol{I})\, \boldsymbol{r} = 2\boldsymbol{D}_l \boldsymbol{s} + \lambda_2 \boldsymbol{p}_{dc}, \tag{18}$$

where $\boldsymbol{D}_l$ is the diagonalization matrix of $\boldsymbol{l}$, the vector $\boldsymbol{p}_{dc}$ is the column-wise vectorization of the tensor $\mathcal{P}(DC)$.

Similar to the solution process for $\boldsymbol{l}$, the solution of $\boldsymbol{r}$ can also be transformed into solving a series of linear equation systems:

$$\boldsymbol{r} = (2(\boldsymbol{D}_l \circ \boldsymbol{D}_l) + \boldsymbol{I})^{-1} (2\boldsymbol{D}_l \boldsymbol{s} + \lambda_2 \boldsymbol{p}_{dc}). \tag{19}$$

This paper utilizes the preconditioned conjugate gradient (PCG) method [BBC*94] to expedite the solution of linear equation systems within the sub-problems of $\mathcal{L}$ and $\mathcal{R}$.

(3) ***Update*** $\boldsymbol{D}$: The corresponding optimization sub-problem for dictionary $\boldsymbol{D}^{k+1}$ is

$$\begin{aligned} \boldsymbol{D}^{k+1} &= \arg\min_{\boldsymbol{D}^T\boldsymbol{D}=\boldsymbol{I}} \left\| \mathcal{R} - \mathcal{P}^{-1}(\boldsymbol{DC}) \right\|_F^2 \\ &= \arg\min_{\boldsymbol{D}^T\boldsymbol{D}=\boldsymbol{I}} \left\| \mathcal{P}\left[\mathcal{R} - \mathcal{P}^{-1}(\boldsymbol{DC}) \right] \right\|_F^2 \\ &= \arg\min_{\boldsymbol{D}^T\boldsymbol{D}=\boldsymbol{I}} \|\mathcal{P}(\mathcal{R}) - \boldsymbol{DC}\|_F^2, \end{aligned} \tag{20}$$

where $\mathcal{P}^{-1}$ is the inverse operation of $\mathcal{P}$, employed to transform tensors into matrices.

According to [BCJ13, LSF*15], the solution to the problem in Eq. (20) is given by

$$\boldsymbol{D}^{k+1} = \boldsymbol{U}\boldsymbol{V}^T, \tag{21}$$

where $\boldsymbol{U}$ and $\boldsymbol{V}$ are obtained through the singular value decomposition of matrix $\mathcal{P}(\mathcal{R})\boldsymbol{C}^T$, that is, $\mathcal{P}(\mathcal{R})\boldsymbol{C}^T = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T$.

(4) ***Update*** $\boldsymbol{C}$: The corresponding optimization sub-problem for sparse code $\boldsymbol{C}^{k+1}$ is

$$\begin{aligned} \boldsymbol{C}^{k+1} &= \arg\min_{\boldsymbol{C}} \frac{\lambda_2}{2} \left\| \mathcal{R} - \mathcal{P}^{-1}(\boldsymbol{DC}) \right\|_F^2 + \lambda_3 \|\boldsymbol{C}\|_0 \\ &= \arg\min_{\boldsymbol{C}} \frac{\lambda_2}{2} \left\| \mathcal{P}\left[\mathcal{R} - \mathcal{P}^{-1}(\boldsymbol{DC}) \right] \right\|_F^2 + \lambda_3 \|\boldsymbol{C}\|_0 \\ &= \arg\min_{\boldsymbol{C}} \frac{\lambda_2}{2} \left\| \mathcal{P}(\mathcal{R}) - \boldsymbol{D}^T\boldsymbol{C} \right\|_F^2 + \lambda_3 \|\boldsymbol{C}\|_0 \\ &= \arg\min_{\boldsymbol{C}} \frac{\lambda_2}{2} \left\| \boldsymbol{D}^T\mathcal{P}(\mathcal{R}) - \boldsymbol{C} \right\|_F^2 + \lambda_3 \|\boldsymbol{C}\|_0. \end{aligned} \tag{22}$$

The problem can be solved by employing the hard thresholding method [BD08], leading to the following solution:

$$\boldsymbol{C}^{k+1} = H_{\sqrt{\frac{2\lambda_3}{\lambda_2}}} \left(\boldsymbol{D}^T \mathcal{P}(\mathcal{R}) \right), \tag{23}$$

where H_{thresh} is the hard thresholding operator:

$$H_{thresh}(t) = \begin{cases} t, & |t| > thresh \\ 0, & otherwise. \end{cases} \tag{24}$$

Subsequently, the reconstructed reflectance component $\hat{\mathcal{R}}$ can be obtained by

$$\hat{\mathcal{R}}^{k+1} = \mathcal{P}^{-1}(\boldsymbol{DC}). \tag{25}$$

(5) ***Convergence criteria***: The algorithm will terminate the iteration when either of the following conditions are satisfied.

$$\begin{cases} \left\| \mathcal{L}^{k+1} - \mathcal{L}^k \right\|_F^2 / \|\mathcal{L}\|_F^2 \le \delta \\ \left\| \mathcal{R}^{k+1} - \mathcal{R}^k \right\|_F^2 / \|\mathcal{R}\|_F^2 \le \delta, \end{cases} \tag{26}$$

where δ is set to 0.001.

3.2. Illuminance Adjustment

After obtaining the illuminance $\mathcal{L}$ and reflectance $\mathcal{R}$ estimates, Gamma correction is adopted to the illuminance as in [CXG*17, LL22] to modify the estimated illuminance. It is important to note that gamma correction is performed in the HSV color space to preserve color information. The Gamma correction of the V channel image $\boldsymbol{L}_v$ of $\mathcal{L}$ is defined as $\boldsymbol{L}'_v = \boldsymbol{L}_v^{\frac{1}{\gamma}}$, where the parameter γ is empirically set as 2.2. The final enhanced result is represented as

$$\hat{\mathcal{S}} = \mathcal{L}' \circ \hat{\mathcal{R}}, \tag{27}$$

where $\mathcal{L}'$ is obtained by converting the gamma-corrected luminance to the RGB color space.

Algorithm 1 outlines the whole process of enhancing low-light images using the proposed BTRetinex model.

Algorithm 1 Low-Light Image Enhancement Based on the BTRetinex Model in Equation (11)

Input: Observed image $\mathcal{S}$, parameters λ_1, λ_2 and λ_3, number of iterations K, stopping criteria δ.

1: Initialize $\mathcal{L}^0 = \mathcal{S}$, $\mathcal{R}^0 = \mathcal{I}$ and $k = 0$;
2: Update $\mathcal{L}^{k+1}$ via Eq. (16);
3: Update $\mathcal{R}^{k+1}$ via Eq. (19);
4: Update D^{k+1} via Eq. (21);
5: Update C^{k+1} via Eq. (23);
6: Check convergence criteria via Eq. (26);
7: If the convergence criteria are not satisfied, set $k = k+1$ and return to the step 2.

Output: The enhanced image $\hat{\mathcal{S}}$ via Eq. (27).

Table 1: *Quantitative comparison of the recent high-performing Retinex-based methods on the MixPG660 dataset.* ***Bold red*** *indicates the best results, while* ***bold green*** *indicates the second-best results.*

Metric	ILNIQE↓	NIQMC↑	VIF↑	Time↓
Input	27.445	4.508	1.000	-
WVM	26.153	4.683	1.315	8.177
Jiep	**25.517**	4.732	1.319	**2.724**
STAR	25.836	4.804	1.212	**3.254**
LR3M	27.909	4.842	0.782	184.302
PnPRetinex	25.998	**4.999**	**1.520**	5.223
BTRetinex(Ours)	**24.819**	**4.939**	**1.328**	41.606

4. Experiments

In this section, we assess the performance of our proposed BTRetinex Model by subjectively and objectively comparing it with state-of-the-art Retinex decomposition and low-light image enhancement methods. We run our experiments using MATLAB 2022b on a PC with AMD Ryzen 5 3600 6-Core Processor and 32GB RAM.

4.1. Implementation Details

Parameters Setting: The empirical parameters λ_1, λ_2 , λ_3 and ε are respectively set as 0.002, 0.005, 10^{-6} and 0.001. The experiments confirm that the proposed model achieves convergence after 8 iterations, demonstrating satisfactory convergence behavior. To strike a balance between performance and algorithm efficiency, we set the maximum number of iterations, K, to 8.

Dataset: We use 660 images collected from MF [FZH*16a], LIME [GLL16], SIRE [WBSS04], DICM [LLK13], NPE [WZHL13], MEF [MZW15] and ExDark [LC19] to test the performance. We temporarily designate the dataset we collect as *MixPG660*, which encompasses low-light images under various lighting conditions and noise levels encountered in real-world scenarios. This dataset effectively serves as a robust benchmark for evaluating the performance of Retinex decomposition and low-light enhancement algorithms.

Evaluation Metrics: We employ two blind image quality assessment metrics, namely Integrated Local Natural Image Quality Evaluator (ILNIQE) [ZZB15] and No-reference Image Quality Metric for Contrast distortion (NIQMC) [GLZ*16], along with one full-reference image quality assessment metric, Visual Information Fidelity (VIF) [SB06], to provide a comprehensive evaluation of the enhanced results. These metrics assess various aspects of image quality, including perceptual quality, contrast and detail preservation. Lower ILNIQE values and higher values of NIQMC and VIF indicate higher image quality.

Methods Included in the Comparison: We compare our proposed BTRetinex model with five state-of-the-art Retinex-based low-light image enhancement methods, including WVM [FZH*16b], Jiep [CXG*17], STAR [XHR*20], Low-Rank Regularized Retinex Model(LR3M) [RYCL20], PnPRetinex [LL22], and the four latest learning-based methods, including ZeroDCE [GLG*20], SCI [MML*22], URetinex [WWZ*22] and LLFlow [WWY*22]. All results of the compared methods are obtained using the parameters recommended by the authors.

It is worth noting that the MSRCR [RJW96], NPE [WZHL13] and LIME [GLL16] methods shown in Figure 1 were proposed relatively early, and they exhibit significant issues in the results of image enhancement, such as color distortion, severe artifacts, excessive brightness enhancement and noise amplification. Therefore, for the sake of conciseness, this paper does not provide a detailed enumeration of the qualitative and quantitative results of these methods.

4.2. Comparing with variation-based Retinex Methods

Figure 2, 3, 4 and Table 1 illustrate the qualitative and quantitative comparison results between our proposed BTRetinex model and the latest state-of-the-art Retinex-based methods, namely LR3M, WVM, JieP, STAR and PnPRetinex, on the *MixPG660* dataset.

Qualitative Comparison: We start our analysis by examining the characteristics of the illuminance and reflectance components.

Regarding the illuminance component, we can draw the following conclusions:

1) Although the WVM method is capable of smoothing out the details of the illuminance component, it exhibits noticeable diffusion effects at strong edges of the illuminance component, leading to a decrease in the accuracy of the reflectance component decomposition. Specifically, this is manifested by a lower grayscale value of the reflectance component at strong edges, resulting in insufficient brightness enhancement and a consequent decrease in the contrast metric NIQMC, as shown in Table 1.
2) The JieP, STAR and LR3M methods all demonstrate effective preservation of the edges in the illuminance component. However, there are still some image details present in the illuminance component, particularly with the STAR and LR3M methods. The illuminance component of these methods lacks sufficient piece-wise smoothness, which, consequently, leads to a decrease in the contrast of the enhanced image.
3) The PnPRetinex method effectively smooths out the details of the illuminance component while preserving strong edges in the image, as demonstrated by the dress in the second row of Figure 2. However, the PnPRetinex method also smoothes out the weak edges in the image, as indicated by the red or yellow boxes in the second row of Figure 2, 3 and 4. This results in enhanced images with higher contrast, as evidenced by the higher value of the contrast evaluation metric, NIQMC. Nevertheless, excessive smoothing of weak edges can lead to uneven brightness and loss of color information in the enhanced image, as observed in the leg area highlighted by the red box in the first row of Figure 2.

Regarding the reflectance component, we can draw the following conclusions:

1) The WVM, JieP, STAR, LR3M and PnPRetinex methods effectively preserve the details and edges of the reflectance component, as demonstrated in the third row of Figures 2, 3 and 4. However, when the input low-light images contain intense noise or artifacts, these methods are unable to effectively remove the

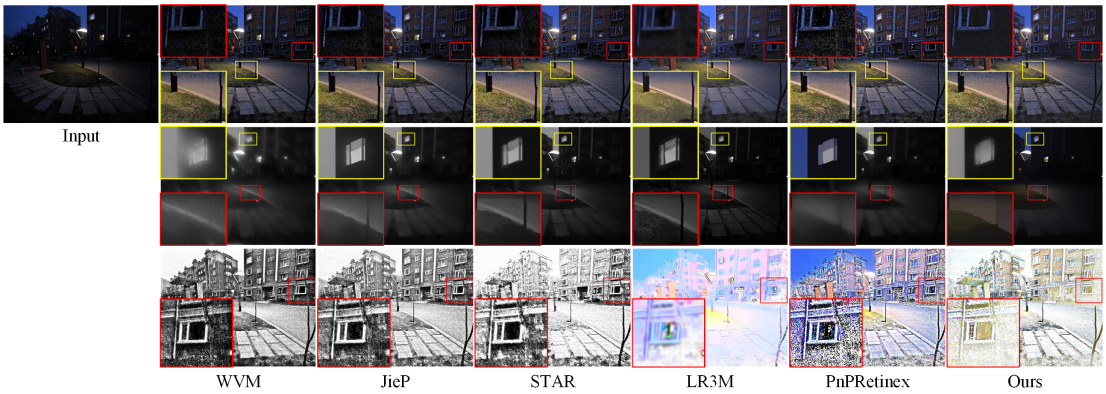

Figure 3: *Comparative analysis of the enhancement and decomposition results between our proposed model and the latest variation-based Retinex methods, namely WVM [FZH*16b], JieP [CXG*17], STAR [XHR*20], LR3M [RYCL20] and PnPRetinex [LL22], under intense noise level. Row 1 represents the enhanced results, while rows 2 and 3 correspond the illuminance and reflectance components, respectively.*

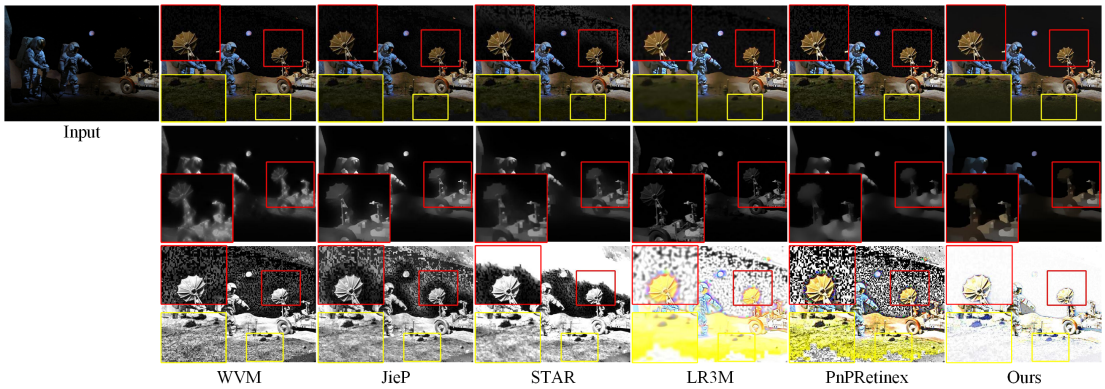

Figure 4: *Comparative analysis of the enhancement and decomposition results between our proposed model and the latest variation-based Retinex methods, namely WVM [FZH*16b], JieP [CXG*17], STAR [XHR*20], LR3M [RYCL20] and PnPRetinex [LL22], under intense artifacts. Row 1 represents the enhanced results, while rows 2 and 3 correspond the illuminance and reflectance components, respectively.*

noise and artifacts from the reflectance component, resulting in noticeable noise and artifacts in the enhanced images, as shown in the first and third row of Figure 3 and 4.

2) The LR3M method can effectively remove noise from the reflectance component. However, at the same time, it results in excessive smoothing of image details, leading to a significant loss of detail information in the enhanced images, as shown in the first row of Figures 2, 3 and 4. This is mainly due to the low-rank constraint imposed on the reflectance component and the equal shrinkage of singular values with different magnitudes. Additionally, in the presence of strong artifacts, the LR3M method fails to completely eliminate artifacts in the reflectance component, as illustrated in the first row of Figure 4.

Based on the results shown in Figures 2, 3 and 4, it can be concluded that our proposed BTRetinex model effectively suppresses noise and artifacts in the enhanced images while preserving image details. This is mainly attributed to the application of a tensor sparse coding prior on the reflectance component, which allows for the suppression of noise and artifacts. Furthermore, our model is capable of generating enhanced images with a piece-wise smooth brightness distribution and higher contrast. This is primarily achieved through the implementation of the modified bilateral total variation constraint on the illuminance component, which ensures both piece-wise smoothness and the preservation of weak structures in the decomposed illuminance component.

Quantitative Comparison: According to Table 1, our model significantly outperforms other methods in terms of ILNIQE, resulting in enhanced images that are comparable to natural images and preserve more details. Moreover, our model ranks second in terms of the NIQMC metric, and its value is very close to that of the PnPRetinex method. This indicates that our model produces enhanced results with excellent contrast. Additionally, our model demonstrates outstanding visual quality, as evidenced by its attainment of the second highest ranking in the VIF metric. However, in terms of efficiency, our method falls behind the WVM, JieP, STAR and PnPRetinex methods, although it still outperforms the LR3M method. Further improvements are necessary to improve the efficiency of our model.

Table 2: *Quantitative comparison of the four latest learning-based methods on the MixPG660 dataset.* ***Bold red*** *indicates the best results, while* ***bold green*** *indicates the second-best results.*

Metric	~~ILNIQE~~↓	~~NIQMC~~↑	VIF↑	Time↓
Input	27.445	4.508	1.000	-
ZeroDCE	25.545	4.930	**1.260**	**0.001**
SCI	26.875	**5.278**	1.229	0.110
URetinex	**24.008**	5.166	1.074	**0.043**
LLFlow	**23.062**	**5.293**	1.171	0.185
BTRetinex(Ours)	24.819	4.939	**1.328**	41.606

4.3. Comparison with Learning-based Methods

Figure 5 and Table 2 illustrate the qualitative and quantitative comparison results between our proposed BTRetinex model and the recent state-of-the-art learning-based methods, namely ZeroDCE, LLFlow, URetinex and SCI, on the MixPG660 dataset.

From Figure 5, it can be observed that, unlike the variation Retinex-based methods discussed in Section 4.2, the learning-based methods exhibit noticeable visual distortions, including color distortions, over-enhancement, and strong artifacts. These distortion issues lead to the instability of the ILNIQE and NIQMC metrics [ZZB15, GLZ*16]. Therefore, in Table 2, we only provide their values without considering them as quantitative indicators. Furthermore, in Figure 5, by comparing with the brighter regions in the input low-light images, it can be observed that our method produces results that closely resemble the visual effects of the input low-light images. Consequently, our model demonstrates superior performance in terms of visual fidelity, as indicated by the Visual Information Fidelity (VIF) metric, as shown in Table 2. In terms of computational efficiency, the learning-based methods exhibit exceptional efficiency due to the utilization of GPU acceleration.

4.4. Limitations

While our proposed BTRetinex model demonstrates strong noise suppression capabilities and visually appealing enhancement results for low-light image enhancement tasks, it does exhibit two primary limitations. Firstly, due to the significant number of image patch operations involved in tensor sparse coding, the BTRetinex model tends to be time-consuming. In the future, we plan to enhance the model's time efficiency and generalization performance by integrating physical imaging models with deep learning techniques. Secondly, for regions in images with high noise levels and low brightness, the BTRetinex model tends to smooth out part structural information while removing noise, leading to a slight deficiency in brightness improvement in these regions. To address this limitation, we intend to introduce a weighted strategy based on image structure information in the tensor sparse coding process.

5. Conclusion

This paper proposed the BTRetinex model as an effective approach for enhancing low-light images by effectively suppressing noise and artifacts while preserving important details. The BTRetinex model employs the modified bilateral total variation technique to efficiently smooth out textures in the illuminance component while preserving weak structures, and utilizes tensor sparse coding to eliminate noise and artifacts from the reflectance component. Experimental results on diverse datasets demonstrate that the proposed method outperforms or performs comparably to state-of-the-art approaches in both qualitative and quantitative evaluations.

Similar to other variation Retinex methods, suitable model parameters need to be manually selected through experimentation, which can be a tedious process. However, once the model parameters are determined, our model exhibits stable and reliable performance. Moreover, while our model demonstrates acceptable computational efficiency, there is still potential for further improvement in terms of optimizing its computational performance.

Acknowledgments

This work was supported by the Guangdong Province Basic and Applied Basic Research Fund Project under Grant 2019A1515011041, Science and Technology Project of Guangzhou under Grant 202103010003, Science and Technology Project in key areas of Foshan under Grant 2020001006285, Xijiang Innovation Team of Zhaoqing under Grant XJCXTD3-2019-04B.

References

[AEB06] AHARON M., ELAD M., BRUCKSTEIN A.: K-svd: An algorithm for designing overcomplete dictionaries for sparse representation. IEEE Transactions on signal processing 54, 11 (2006), 4311–4322. 4

[BBC*94] BARRETT R., BERRY M., CHAN T. F., DEMMEL J., DONATO J., DONGARRA J., EIJKHOUT V., POZO R., ROMINE C., VAN DER VORST H.: Templates for the solution of linear systems: building blocks for iterative methods. SIAM, 1994. 6

[BCJ13] BAO C., CAI J.-F., JI H.: Fast sparsity-based orthogonal dictionary learning for image restoration. In Proceedings of the IEEE International Conference on Computer Vision (2013), pp. 3384–3391. 5, 6

[BD08] BLUMENSATH T., DAVIES M. E.: Iterative thresholding for sparse approximations. Journal of Fourier analysis and Applications 14 (2008), 629–654. 6

[BW86] BRAINARD D. H., WANDELL B. A.: Analysis of the retinex theory of color vision. JOSA A 3, 10 (1986), 1651–1661. 1

[CJSY14] CAI J.-F., JI H., SHEN Z., YE G.-B.: Data-driven tight frame construction and image denoising. Applied and Computational Harmonic Analysis 37, 1 (2014), 89–105. 5

[CLKL14a] CHO H., LEE H., KANG H., LEE S.: Bilateral texture filtering. ACM Transactions on Graphics (TOG) 33, 4 (2014), 1–8. 3

[CLKL14b] CHO H., LEE H., KANG H., LEE S.: Bilateral texture filtering. ACM Transactions on Graphics (TOG) 33, 4 (2014), 1–8. 4

[CXG*17] CAI B., XU X., GUO K., JIA K., HU B., TAO D.: A joint intrinsic-extrinsic prior model for retinex. In Proceedings of the IEEE international conference on computer vision (2017), pp. 4000–4009. 1, 2, 3, 4, 6, 7, 8

[CYZ*20] CHANG Y., YAN L., ZHAO X.-L., FANG H., ZHANG Z., ZHONG S.: Weighted low-rank tensor recovery for hyperspectral image restoration. IEEE transactions on cybernetics 50, 11 (2020), 4558–4572. 2, 5

[DHQ*18] DU Y., HAN G., QUAN Y., YU Z., WONG H.-S., CHEN C. P., ZHANG J.: Exploiting global low-rank structure and local sparsity nature for tensor completion. IEEE Transactions on Cybernetics 49, 11 (2018), 3898–3910. 5

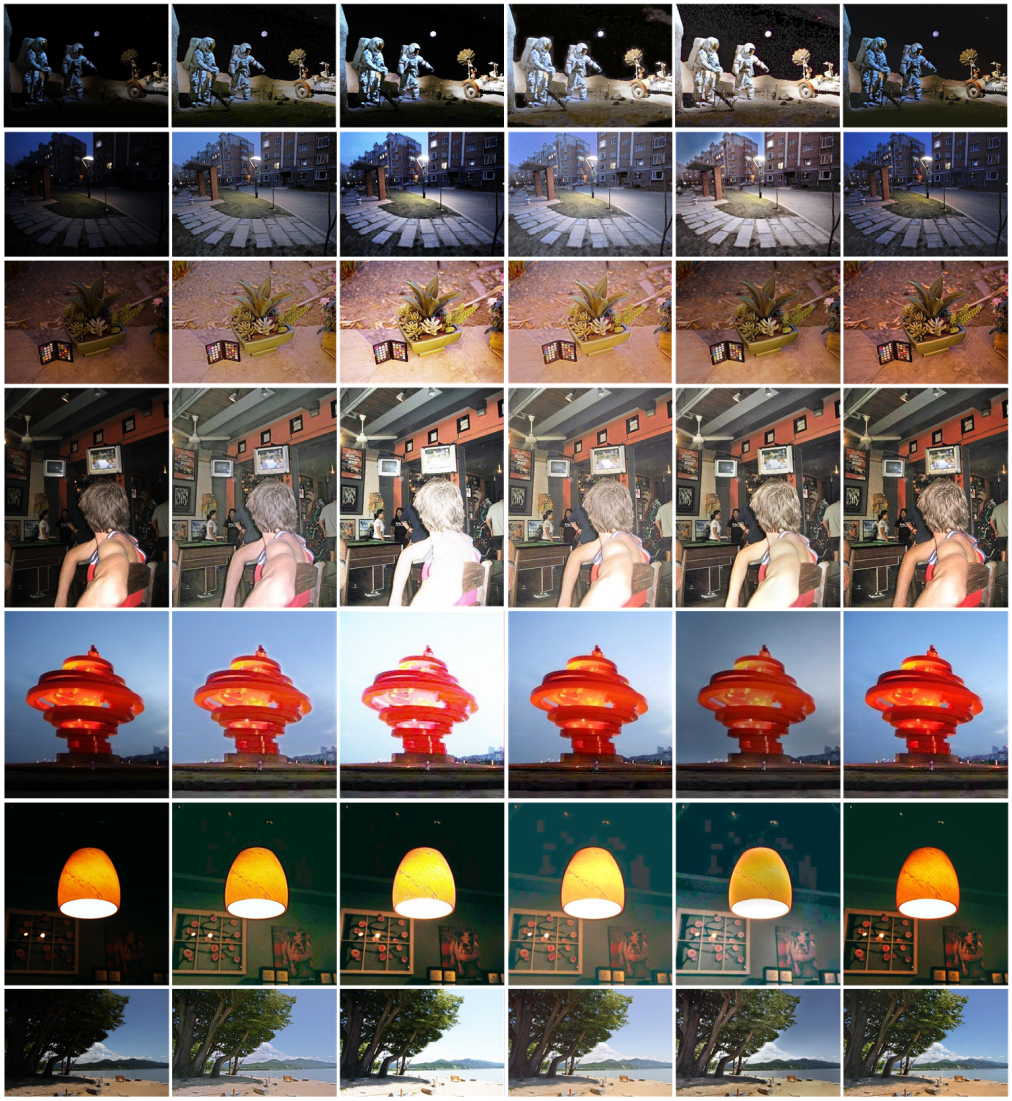

Figure 5: *Comparative analysis of the enhancement results between our proposed model and the recent learning-based methods, namely ZeroDCE [GLG*20], SCI [MML*22], URetinex [WWZ*22] and LLFlow [WWY*22], across various scenarios.*

[DWL*12] DUAN G., WANG H., LIU Z., DENG J., CHEN Y.-W.: K-cpd: Learning of overcomplete dictionaries for tensor sparse coding. In Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012) (2012), IEEE, pp. 493–496. 5

[EA06] ELAD M., AHARON M.: Image denoising via sparse and redundant representations over learned dictionaries. IEEE Transactions on Image processing 15, 12 (2006), 3736–3745. 4

[FLZ*15] FU X., LIAO Y., ZENG D., HUANG Y., ZHANG X.-P., DING X.: A probabilistic method for image enhancement with simultaneous illumination and reflectance estimation. IEEE Transactions on Image Processing 24, 12 (2015), 4965–4977. 2, 3

[FZH*16a] FU X., ZENG D., HUANG Y., LIAO Y., DING X., PAISLEY J.: A fusion-based enhancing method for weakly illuminated images. Signal Processing 129 (2016), 82–96. 7

[FZH*16b] FU X., ZENG D., HUANG Y., ZHANG X.-P., DING X.: A weighted variational model for simultaneous reflectance and illumination estimation. In Proceedings of the IEEE conference on computer vision and pattern recognition (2016), pp. 2782–2790. 1, 2, 4, 7, 8

[GLG*20] GUO C., LI C., GUO J., LOY C. C., HOU J., KWONG S., CONG R.: Zero-reference deep curve estimation for low-light image enhancement. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (2020), pp. 1780–1789. 2, 7, 10

[GLL16] GUO X., LI Y., LING H.: Lime: Low-light image enhancement via illumination map estimation. IEEE Transactions on image processing 26, 2 (2016), 982–993. 2, 7

[GLZ*16] GU K., LIN W., ZHAI G., YANG X., ZHANG W., CHEN C. W.: No-reference quality metric of contrast-distorted images based on information maximization. IEEE transactions on cybernetics 47, 12 (2016), 4559–4565. 7, 9

[HST12] HE K., SUN J., TANG X.: Guided image filtering. IEEE

transactions on pattern analysis and machine intelligence 35, 6 (2012), 1397–1409. 2

[HXXC22] He L., Xie Y., Xie S., Chen Z.: Structure-preserving texture smoothing via scale-aware bilateral total variation. IEEE Transactions on Circuits and Systems for Video Technology (2022). 3, 4, 5

[JRW97a] Jobson D. J., Rahman Z.-u., Woodell G. A.: A multiscale retinex for bridging the gap between color images and the human observation of scenes. IEEE Transactions on Image processing 6, 7 (1997), 965–976. 1

[JRW97b] Jobson D. J., Rahman Z.-u., Woodell G. A.: Properties and performance of a center/surround retinex. IEEE transactions on image processing 6, 3 (1997), 451–462. 1

[Lan77] Land E. H.: The retinex theory of color vision. Scientific american 237, 6 (1977), 108–129. 1

[Lan83] Land E. H.: Recent advances in retinex theory and some implications for cortical computations: color vision and the natural image. Proceedings of the National Academy of Sciences 80, 16 (1983), 5163–5169. 1

[LC19] Loh Y. P., Chan C. S.: Getting to know low-light images with the exclusively dark dataset. Computer Vision and Image Understanding 178 (2019), 30–42. 7

[LL22] Lin Y.-H., Lu Y.-C.: Low-light enhancement using a plug-and-play retinex model with shrinkage mapping for illumination estimation. IEEE Transactions on Image Processing 31 (2022), 4897–4908. 1, 2, 3, 4, 6, 7, 8

[LLK13] Lee C., Lee C., Kim C.-S.: Contrast enhancement based on layered difference representation of 2d histograms. IEEE transactions on image processing 22, 12 (2013), 5372–5384. 7

[LLY*18] Li M., Liu J., Yang W., Sun X., Guo Z.: Structure-revealing low-light image enhancement via robust retinex model. IEEE Transactions on Image Processing 27, 6 (2018), 2828–2841. 2

[LSF*15] Liu Y., Shang F., Fan W., Cheng J., Cheng H.: Generalized higher order orthogonal iteration for tensor learning and decomposition. IEEE transactions on neural networks and learning systems 27, 12 (2015), 2551–2563. 6

[MML*22] Ma L., Ma T., Liu R., Fan X., Luo Z.: Toward fast, flexible, and robust low-light image enhancement. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (2022), pp. 5637–5646. 2, 7, 10

[MMOC11] Ma W., Morel J.-M., Osher S., Chien A.: An l 1-based variational model for retinex theory and its application to medical images. In CVPR 2011 (2011), IEEE, pp. 153–160. 1

[MPS10] Morel J. M., Petro A. B., Sbert C.: A pde formalization of retinex theory. IEEE Transactions on Image Processing 19, 11 (2010), 2825–2837. 1

[MZW15] Ma K., Zeng K., Wang Z.: Perceptual quality assessment for multi-exposure image fusion. IEEE Transactions on Image Processing 24, 11 (2015), 3345–3356. 7

[OF97] Olshausen B. A., Field D. J.: Sparse coding with an overcomplete basis set: A strategy employed by v1? Vision research 37, 23 (1997), 3311–3325. 4

[PDCRM05] Provenzi E., De Carli L., Rizzi A., Marini D.: Mathematical definition and analysis of the retinex algorithm. JOSA A 22, 12 (2005), 2613–2621. 1

[QHJ15] Quan Y., Huang Y., Ji H.: Dynamic texture recognition via orthogonal tensor dictionary learning. In Proceedings of the IEEE international conference on computer vision (2015), pp. 73–81. 5

[RJW96] Rahman Z.-u., Jobson D. J., Woodell G. A.: Multi-scale retinex for color image enhancement. In Proceedings of 3rd IEEE international conference on image processing (1996), vol. 3, IEEE, pp. 1003–1006. 1, 2, 3, 7

[ROF92] Rudin L. I., Osher S., Fatemi E.: Nonlinear total variation based noise removal algorithms. Physica D: nonlinear phenomena 60, 1-4 (1992), 259–268. 3

[RYCL20] Ren X., Yang W., Cheng W.-H., Liu J.: Lr3m: Robust low-light enhancement via low-rank regularized retinex model. IEEE Transactions on Image Processing 29 (2020), 5862–5876. 2, 3, 4, 7, 8

[SB06] Sheikh H. R., Bovik A. C.: Image information and visual quality. IEEE Transactions on image processing 15, 2 (2006), 430–444. 7

[WBSS04] Wang Z., Bovik A. C., Sheikh H. R., Simoncelli E. P.: Image quality assessment: from error visibility to structural similarity. IEEE transactions on image processing 13, 4 (2004), 600–612. 7

[WCZ*19] Wang Y., Cao Y., Zha Z.-J., Zhang J., Xiong Z., Zhang W., Wu F.: Progressive retinex: Mutually reinforced illumination-noise perception network for low-light image enhancement. In Proceedings of the 27th ACM international conference on multimedia (2019), pp. 2015–2023. 2

[WWY*22] Wang Y., Wan R., Yang W., Li H., Chau L.-P., Kot A.: Low-light image enhancement with normalizing flow. In Proceedings of the AAAI Conference on Artificial Intelligence (2022), vol. 36, pp. 2604–2612. 2, 7, 10

[WWYL18] Wei C., Wang W., Yang W., Liu J.: Deep retinex decomposition for low-light enhancement. arXiv preprint arXiv:1808.04560 (2018). 2

[WWZ*22] Wu W., Weng J., Zhang P., Wang X., Yang W., Jiang J.: Uretinex-net: Retinex-based deep unfolding network for low-light image enhancement. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (2022), pp. 5901–5910. 2, 7, 10

[WZHL13] Wang S., Zheng J., Hu H.-M., Li B.: Naturalness preserved enhancement algorithm for non-uniform illumination images. IEEE transactions on image processing 22, 9 (2013), 3538–3548. 1, 2, 7

[XHR*20] Xu J., Hou Y., Ren D., Liu L., Zhu F., Yu M., Wang H., Shao L.: Star: A structure and texture aware retinex model. IEEE Transactions on Image Processing 29 (2020), 5022–5037. 1, 2, 3, 4, 7, 8

[XYXJ12] Xu L., Yan Q., Xia Y., Jia J.: Structure extraction from texture via relative total variation. ACM transactions on graphics (TOG) 31, 6 (2012), 1–10. 3, 5

[ZGM*21] Zhang Y., Guo X., Ma J., Liu W., Zhang J.: Beyond brightening low-light images. International Journal of Computer Vision 129 (2021), 1013–1037. 2

[ZZB15] Zhang L., Zhang L., Bovik A. C.: A feature-enriched completely blind image quality evaluator. IEEE Transactions on Image Processing 24, 8 (2015), 2579–2591. 7, 9

[ZZG19] Zhang Y., Zhang J., Guo X.: Kindling the darkness: A practical low-light image enhancer. In Proceedings of the 27th ACM international conference on multimedia (2019), pp. 1632–1640. 2