

LR3M: Robust Low-Light Enhancement via Low-Rank Regularized Retinex Model

Xutong Ren, Wenhan Yang, *Member, IEEE*,
 Wen-Huang Cheng *Senior Member, IEEE*, Jiaying Liu, *Senior Member, IEEE*

Abstract—Noise causes unpleasant visual effects in low-light image/video enhancement. In this paper, we aim to make the enhancement model and method aware of noise in the whole process. To deal with heavy noise which is not handled in previous methods, we introduce a robust low-light enhancement approach, aiming at well enhancing low-light images/videos and suppressing intensive noise jointly. Our method is based on the proposed Low-Rank Regularized Retinex Model (LR3M), which is the first to inject low-rank prior into a Retinex decomposition process to suppress noise in the reflectance map. Our method estimates a piece-wise smoothed illumination and a noise-suppressed reflectance sequentially, avoiding remaining noise in the illumination and reflectance maps which are usually presented in alternative decomposition methods. After getting the estimated illumination and reflectance, we adjust the illumination layer and generate our enhancement result. Furthermore, we apply our LR3M to video low-light enhancement. We consider inter-frame coherence of illumination maps and find similar patches through reflectance maps of successive frames to form the low-rank prior to make use of temporal correspondence. Our method performs well for a wide variety of images and videos, and achieves better quality both in enhancing and denoising, compared with the state-of-the-art methods.

Index Terms—Low-Light Enhancement, Denoising, Retinex Model, Low-Rank Decomposition

I. INTRODUCTION

WITH the booming of social media such as Facebook and YouTube, it becomes popular to take photograph or video to record people's daily lives and share them with others through the Internet. However, due to lacking the professional shooting skill for most users, many photos are captured under low-light circumstance due to backlight, under-exposure and dark environment. These photos, called low-light images, suffer from low visibility, low contrast and noise. It is desirable to directly enhance the images and apply denoising during the camera processing pipeline [1], taking the raw

Copyright (c) 2018 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending an email to pubs-permissions@ieee.org.

This work was supported by National Natural Science Foundation of China under contract No.61772043, Beijing Natural Science Foundation under contract No.L182002, and the National Key R&D Program of China under Grand No.2018AAA0102700. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Jana Ehmann. (Corresponding author: Jiaying Liu.)

Xutong Ren, Wenhan Yang, Jiaying Liu are with Institute of Computer Science and Technology, Peking University, Beijing, 100080, China, e-mail: {tonghelen, yangwenhan, liujiaying}@pku.edu.cn.

Wen-Huang Cheng is with National Chiao Tung University, Hsinchu, 30010, Taiwan, e-mail: whcheng@nctu.edu.tw.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

signal as input, where the dynamic of pixel values (in raw Bayer image) is higher. However, in many cases, it might be not easy to access that pipeline for hardware and economical reasons. Therefore, many researches focus on post-processing techniques to enhance the visual quality of these images. In this paper, we follow the second route.

There are several kinds of methods to enhance low-light images, including histogram equalization [2, 3], inverse domain operation [4, 5], Retinex decomposition [6–9], and deep learning [10, 11]. Retinex decomposition-based methods consider the scene in human's eyes as the product of reflectance and illumination layers. The enhanced results are produced by adjusting the corresponding layers. The earliest methods directly regard the decomposed reflectance layer as the enhanced result [6–9]. The single-scale Retinex (SSR) [6] and multi-scale Retinex (MSR) [7] utilize the Gaussian filter to build Retinex representation. In [9], bilateral filter is utilized to remove halo artifacts. Later approaches adjust both illumination and reflectance layers, and reconstruct the enhanced result by combining them. In [12–14], variational models estimate the piecewise continuous reflectance layer and smooth illumination layer of the Retinex model.

In these methods, the logarithmic transformation is widely used because of the ease of modeling by most Retinex based methods. However, the work [15] found that logarithmic transformation is not a good choice. To cover the shortage of logarithmic form, a weighted variational model is provided to estimate both the reflectance and the illumination. The reconstructed results are visually pleasing in light distribution but contain observable noise, which impairs the overall visual quality. Another work [16] pays attention to strengthening illumination map by imposing a structure prior. It effectively enhances the contrast. However, it leads to over-exposure and detail loss in bright regions. Besides, since the unprocessed reflectance contains much noise, after enhancement, noise will be amplified. An ad-hoc denoising procedure via BM3D [17] is applied. However, there are remaining noise or over-smoothed details. Yue *et al.* [18] concentrated on intrinsic image decomposition and introduces constraints on both reflectance and illumination layers. They do not take noise as an influence factor in the decomposition procedure.

Recently, Li *et al.* [19] proposed a structure-revealing low-light enhancement approach based on a Refined Retinex model. They tried to remove noise by suppressing small gradients in the reflectance component. However, adjusting the gradients becomes less useful when there is heavy noise. Furthermore, their model assumes that the illumination component is smooth

enough, which leads to observable noise appearing in the reflectance map. But using an alternating direction minimization algorithm makes the left noise disturb the estimation of both illumination and reflectance maps.

In this paper, we consider noise as a non-negligible factor in Retinex based decomposition. We aim to make the proposed model and method aware of noise in the whole process, not in the form of a separate ad-hoc operation. Compared to previous methods that only consider light noise modeling, we also aim to model and remove heavy noise in the low-light image. Specifically, we incorporate low-rank minimization into Retinex decomposition. A robust low-light enhancement method via **Low-Rank Regularized Retinex Model** (LR3M) is proposed to simultaneously enhance the images and remove noise. Considering that, estimating the illumination and reflectance simultaneously but alternatively causes remaining noise in both illumination and reflectance maps, our method chooses to estimate both the illumination and the reflectance maps in a sequential manner. That is, we first estimate the piece-wise illumination map, independent from the reflectance map. Then, we refine the reflectance based on both the refined illumination and the original image. In this step, low rank prior is enforced on the reflectance map to suppress noise. After obtaining the preferred illumination and reflectance, the final enhancement result is generated by combining the reflectance and the Gamma corrected illumination.

This paper is an extension of our previous conference paper [20]. Based on the sequential decomposition in the preliminary work, we additionally inject a low-rank prior to regularize the estimation of the reflectance map and suppress noise in this map. Moreover, we apply our LR3M to robust video low-light enhancement, with inter-frame illumination coherence constraint and more structural correspondence from adjacent frames. We also add extensive experimental analysis to evaluate the effectiveness of the proposed framework on both images and videos. In summary, the contributions of this paper lie in three aspects:

- We propose a robust low-light enhancement method that can both enhance the low-light image/video and denoise jointly in a sequential manner. We demonstrate that an alternative Retinex decomposition causes remaining noise in both illumination and reflectance maps, which impairs overall visual quality. Therefore, our LR3M method is based on a sequential decomposition to make the illumination map free of noise.
- We are the first to inject the low-rank prior into the Retinex decomposition method. Noise is considered as an explicit term in a robust Retinex model. Then, different priors are enforced to estimate illumination and reflectance maps. The reflectance map is constrained by low-rank prior, which facilitates noise removal in both the reflectance map and final enhanced result.
- Our enhancement method can be applied to both images and videos. We consider the inter-frame coherence of illumination maps and find matched blocks through reflectance images of successive frames so as to maintain temporal consistency. In video enhancement case, our low-rank prior is enforced on the three-dimensional pixel

volume flattened from the combination of temporally neighboring patches, which considers the temporal context and benefits the temporal consistency. Extensive experimental results demonstrate the effectiveness of the proposed method objectively and subjectively.

The rest of the paper is organized as follows: Sec. II introduces related Retinex decomposition based low-light image enhancement methods and low-rank minimization method. In Sec. III, the proposed approach is introduced and analyzed, and the algorithm is described. In addition, we elaborate our extension method to handle video low light enhancement. Experimental results are shown in Sec. IV and concluding remarks are given in Sec. V.

II. RELATED WORKS

A. Low-Light Image Enhancement

Low-light image enhancement approaches amplify illumination and improve visibility of dark images. They are classified into four categories: histogram equalization-based, inverse domain-based, Retinex decomposition-based, and deep learning-based.

The most intuitive and simplest way is to directly amplify the illumination of a low-light image. However, this operation results in over-saturation and detail loss in bright areas. *Histogram equalization (HE) based methods* flatten the histogram and stretch the dynamic range of the intensity, alleviating the above problems. To adapt to various intensity distributions of regions rather than only considering global histogram of the entire image, a variant of HE, contrast limited adaptive histogram equalization (CLAHE) [3] is designed to perform HE in each divided block with contrast limiting to mitigate over-enhanced details. A dynamic HE method [2] is developed to perform contrast enhancement on low-light images. Later on, edge-aware tone mapping methods such as edge-preserving decomposition [21] and local Laplacian filtering [22] are proposed to alleviate block artifacts. Recently, a generalized Gaussian mixture-based fusion method [23] is proposed to restore unevenly illuminated images from under-exposure, normal-exposure, and over-exposure images. A dual-exposure fusion algorithm [24] using illumination estimation techniques is proposed to provide an accurate contrast and lightness enhancement. Without a special consideration on noise, the noise and artifacts will be amplified in their results. Some researchers [4, 5] noticed the similarity between haze images and the inverted low-light images. Thus, these *inverse domain-based methods* applied dehazing methods to enhance low-light image.

To jointly adjust illumination and suppress noise, *Retinex theory-based methods* are put forward. The theory assumes that, images are decomposed into reflectance and illumination layers. Then, the enhanced results are reconstructed from the corresponding adjusted layers. The earliest methods utilize the decomposed reflectance layer as the enhanced result [6–9]. The single-scale Retinex (SSR) [6] and multi-scale Retinex (MSR) [7] utilize the Gaussian filter, while [9] uses the bilateral filter to remove halo artifacts. Successive methods manipulate both illumination and reflectance layers to produce the enhanced

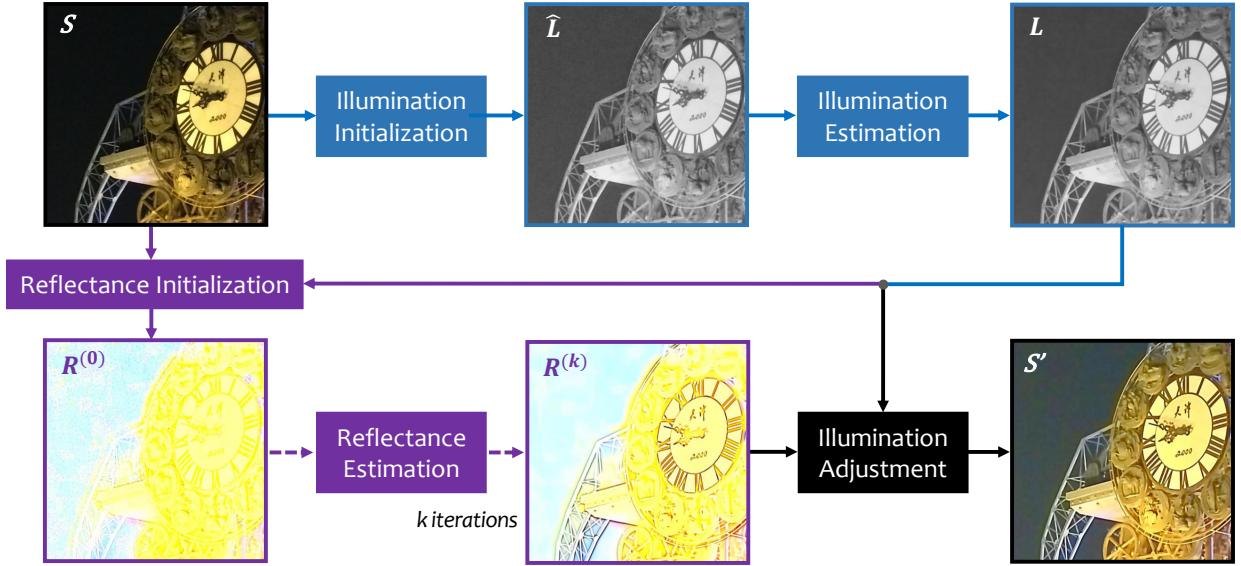


Fig. 1: Framework of the proposed method. We first use an initial illumination to refine the final illumination. Following that we use the obtained illumination map and the input image S to get an initial reflectance map $R^{(0)}$. Then we estimate a noise-free reflectance map based on low-rank regulation. Finally the enhanced image is generated from the illumination and reflectance after the illumination adjustment.

result. In [12–14], variational models are used to estimate the smooth illumination layer and piecewise continuous reflectance layer of the Retinex model. Li *et al.* [19] built a robust Retinex model with an explicit injected noise term and made the first attempt to estimate a noise map out of that model. These methods consider less on the constraints on the reflectance layer, and the latent intensive noises in the low-light regions are usually amplified.

Deep learning-based low-light image enhancement methods have also been studied. Yang *et al.* [10] proposed to enhance low-light images by coupled dictionary learning. Lore *et al.* [11] used a deep auto-encoder named Low-Light Net (LLNet) to perform contrast enhancement and denoising. In [25], deeply root in multi-scale Retinex representation, a feed-forward convolutional neural network with different Gaussian convolution kernels is proposed to learn an end-to-end mapping between dark and bright images. In [26], Wang *et al.* proposes a deep Retinex-Net including a DecomNet for decomposition and an Enhance-Net for illumination adjustment. In [27], Yang *et al.* make the attempt in the semi-supervised learning for low-light image enhancement. In this work, a deep recursive band representation is built to connect fully-supervised and un-supervised learning frameworks and integrate their superiorities. The performance of these works rely heavily on the quality of datasets. Due to the lack of a good metric to evaluate various aspects of the overall quality of the enhanced results, *e.g.* detail preservation, visual naturalness and contrast distribution, their results are not satisfying in some visual aspects.

B. Low-Rank Minimization

Low-rank minimization is a commonly used tool in image completion and denoising. The intrinsic content of a matrix is

usually assumed to be low-rank, because noise is distributed more randomly than the principal component of the matrix. This method recovers or completes a degraded matrix by minimizing the rank of an input corrupted matrix. Given an input matrix P , the original low-rank minimization problem is presented as follows:

$$\min \text{rank}(L), \quad \text{s. t. } L_\Omega = P_\Omega, \quad (1)$$

where Ω represents locations of known elements.

The original low-rank minimization problem in Eqn. (1) is NP-hard, and cannot be solved efficiently. In [28], the problem is relaxed by replacing rank constraint with nuclear norm of the matrix as follows:

$$\min \|L\|_*, \quad \text{s. t. } L_\Omega = P_\Omega, \quad (2)$$

where $\|L\|_*$ is the nuclear norm of matrix L , the sum of the singular values. Eqn. (2) is equivalent to:

$$\operatorname{argmin}_L \|L - P\|_F^2 + \lambda \|L\|_*. \quad (3)$$

Wright *et al.* [29] introduced the iterative thresholding approach to solve a relaxed convex form of the problem. The accelerated proximal gradient approach is provided in [29, 30]. Zhang *et al.* further utilized the truncated nuclear norm [31, 32] to minimize the sum of small singular values. Ono *et al.* [33] proposed the block nuclear norm, which provides a superior characterization of the texture component.

Some works apply low-rank minimization to image completion and denoising. In [34], a spatially adaptive iterative singular-value thresholding method is developed to restore images. In [35, 36], low-rank prior is combined with multi-planar autoregressive model for image completion. In our paper, we are the first to introduce low-rank prior into a

Retinex model for low-light image/video enhancement, suppressing noise while enhancing the illumination. We adopt the augmented Lagrange multiplier minimization algorithm and iterative singular-value thresholding method [34] to solve this joint problem.

III. ROBUST LOW-LIGHT ENHANCEMENT

In this section, we construct a low-rank constrained optimization function to estimate the illumination and reflectance maps of the Retinex model. Then, a sequential solution is provided for robust low light image enhancement. Finally, the proposed method is transferred into video enhancement. Fig. 1 shows the framework of our method.

A. Design Methodology

The traditional Retinex model considers an image S as the multiplication of a reflectance layer R and an illumination layer L physically:

$$S = R \circ L, \quad (4)$$

where the operator \circ denotes the element-wise multiplication.

Reflectance R depicts the intrinsic property of captured objects, which is considered to be consistent under any lightness conditions and full of structural details. The illumination L represents the various lightness on objects. It is piecewise continuous and preserves major edges without small gradients [37, 38].

Knowing that low light may introduce much noise to the image and enhance the picture inevitably intensifies the noise at the same time, we consider a robust Retinex model [19] with an additional noise term N as follows:

$$S = R \circ L + N. \quad (5)$$

Many methods focus on the illumination component L and simply take $R' = S/L$ as the obtained reflectance, which actually keeps most unpleasant noise intact in the reflectance image for $R' = R + N/L$. Thus, those methods always lead to noisy results and often require an extra denoising procedure. In this case, two properties are preferred in a Retinex-based low-light enhancement method:

- *Full Noise Awareness.* Noise has a negative effect on visual quality of enhanced results and is also a non-negligible factor in low-light enhancement. Most previous methods suppress noise by pre/post-processing, which easily leads to residual noises or over-smoothed details in the results, as shown in Fig. 2. Thus, a desirable low light enhancement method should be fully aware of noise and handle it adaptively through the whole enhancement process.
- *Sequential Optimization.* Some methods [15, 19, 39] use term $\|R \cdot L - S\|_F^2$ in their equations, and in order to calculate both R and L simultaneously, they iteratively update each variable while regarding the other variables as constants. In other words, L is calculated on the basis of the previous result of R in every iteration and R is estimated in a same way. However, during these procedures, the noise often observed in the reflectance

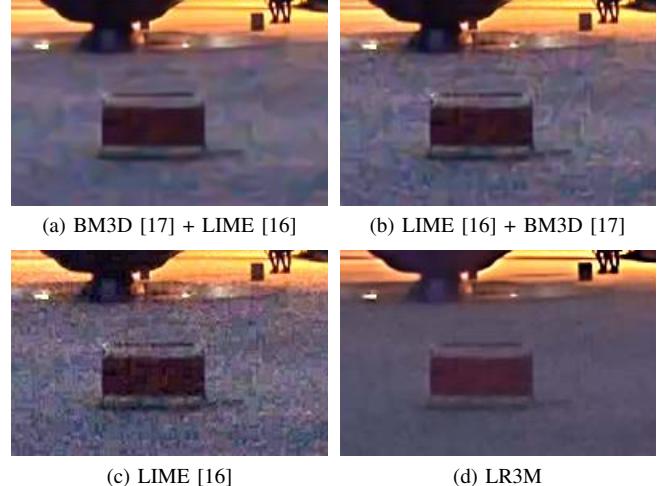


Fig. 2: Visual results of our method and methods using denoising as pre/post-processing methods. (a) and (b): results using low-light enhancement and denoising methods in different orders. (c): result of LIME [16]. (d): result of LR3M.

image continuously impairs the expected illumination map L , which leads to a more degraded reflectance map R in following iterations. This implicates the potential superiority of a one-round sequential method to an alternative updating one. The examples of two different methods can be found in Sec. IV-B.

Based on this consideration, in our work, we estimate the illumination and the reflectance in a successive sequence so that they can be more easily separated compared to an iterative manner. Besides, the low-rank prior is injected into an optimization function to suppress noise in the reflectance layer during the whole optimization process. Thus, our method is by nature capable of adjusting global illumination and removing noise. In the next section, we give the overall objective function and elaborate our proposed method.

B. Single Image Low-Light Enhancement

For low-light images, it usually suffers from darkness and unbalanced illumination distributions. Estimating R and L from S is an ill-posed problems, thus we construct an optimization function that solves the inverse problem of Eqn. (4) with regularization terms:

$$\begin{aligned} \operatorname{argmin}_{L,R} & \|R \circ L - S\|_F^2 + \alpha \|\nabla L\|_1 + \beta \|\nabla R - G\|_F^2 \\ & + \omega \sum_i \|NN_i(R)\|_*, \end{aligned} \quad (6)$$

where α , β and ω are weighting parameters that control the importance of different terms. $\|\cdot\|_F$, $\|\cdot\|_1$ and $\|\cdot\|_*$ represent the Frobenius norm, ℓ_1 norm and nuclear norm, respectively.

In addition, ∇ is the first order differential operator. $NN_i(\cdot)$ is a patch extraction operation that collects similar patches of the patch located at i . That is, $NN_i(R) = [R_{i_1}, R_{i_2}, \dots, R_{i_k}] \in \mathbb{R}^{b^2 \times k}$ is the similar patch group of the reference patch R_{i_1} , where b is the patch size and k is the number of nearest neighbors of the reference patch. G is the adjusted gradient

of the observed image S . The role of each term in Eqn. (6) is interpreted below:

- $\|R \circ L - S\|_F^2$ constrains the fidelity between the observed image S and the recomposed one $R \circ L$.
- $\|\nabla L\|_1$ enforces the total variation sparsity and piecewise smoothness of the illumination map L .
- $\|NN_i(R)\|_*^2$ minimizes the rank of matrices of similar patches of R , so as to make it get rid of noise and artifacts.
- $\|\nabla R - G\|_F^2$ minimizes the distance between the gradient of the reflectance R and that of the observed image S (estimated by G), so that the contrast of the final result can be strengthened. To get a preferable reflectance estimation R whose gradients are suppressed in smooth areas and fully kept at edge regions, G is inferred from an adjusted ∇S [19, 20] as follows:

$$G = (1 + \lambda e^{-|\nabla \hat{S}|/\sigma}) \circ \nabla \hat{S}, \quad (7)$$

$$\nabla \hat{S} = \begin{cases} 0, & \text{if } |\nabla S| < \varepsilon, \\ \nabla S, & \text{otherwise,} \end{cases} \quad (8)$$

where λ controls the degree of the amplification, σ controls the amplification rate of different gradients and ε is the threshold that filters small gradients. By suppressing small gradients first, this equation minimizes the possible noise and then strengthens the overall gradients with adaptive proportions.

In the following, we propose sequential solution to estimate the undisturbed illumination L using an initial illumination map \hat{L} and the preferred reflectance R . For each observed image, \hat{L} and G only need to be calculated once.

1) *Solution to L-Problem (Illumination Estimation):* Collecting all terms related to L in Eqn. (6), we have:

$$\operatorname{argmin}_L \|R \circ L - S\|_F^2 + \alpha \|\nabla L\|_1. \quad (9)$$

The ℓ_1 norm together with the gradient operation makes it complicated and time consuming to solve the problem. Therefore, we estimate L from an initial estimation \hat{L} :

$$\operatorname{argmin}_L \|L - \hat{L}\|_F^2 + \alpha \|\nabla L\|_1. \quad (10)$$

Same as LIME [16], we assume that, for color images, three channels share the same illumination map. Therefore, the initial illumination map \hat{L} is estimated as the average of the input image in RGB color space:

$$\hat{L}(x) = \frac{1}{3} \sum_{c \in \{R, G, B\}} S^c(x). \quad (11)$$

Furthermore, we find the relationship below holds true:

$$\lim_{\epsilon \rightarrow 0^+} \sum_x \sum_{d \in \{h, v\}} \frac{(\nabla_d L(x))^2}{|\nabla_d \hat{L}(x)| + \epsilon} = \|\nabla L\|_1. \quad (12)$$

Thus, we use the alternative $\sum_x \sum_{d \in \{h, v\}} \frac{(\nabla_d L(x))^2}{|\nabla_d \hat{L}(x)| + \epsilon}$ to approximate $\|\nabla L\|_1$. As a result, the approximate problem to Eqn. (10) can be written as follows:

$$\operatorname{argmin}_L \|L - \hat{L}\|_F^2 + \alpha \sum_x \sum_{d \in \{h, v\}} \frac{(\nabla_d L(x))^2}{|\nabla_d \hat{L}(x)| + \epsilon}. \quad (13)$$

This change does not influence the result much because according to the first term $\|L - \hat{L}\|_F^2$, the gradients of L should also be similar to those of \hat{L} . For convenience, we put Eqn. (13) in a simpler form, where $A_d(x)$ represents $\frac{\alpha}{|\nabla_d \hat{L}(x)| + \epsilon}$:

$$\operatorname{argmin}_L \|L - \hat{L}\|_F^2 + \sum_x \sum_{d \in \{h, v\}} A_d(x) \cdot (\nabla_d L(x))^2. \quad (14)$$

As can be observed, Eqn. (14) only involves quadratic terms. Thus, after differentiating Eqn. (14) with respect to L and setting the derivative to 0, the problem can be directly figured out by solving the following:

$$\left(I + \sum_{d \in \{h, v\}} D_d^T \operatorname{Diag}(a_d) D_d \right) l = \hat{l}, \quad (15)$$

where I is the identity matrix with proper size and D contains D_h and D_v , which are the Toeplitz matrices from the discrete gradient operators with forward difference. Furthermore, x is the vectorized version of X and the operator $\operatorname{Diag}(x)$ is to construct a diagonal matrix using vector x . Then, we can easily solve it to obtain the evaluated L :

$$l = \left(I + \sum_{d \in \{h, v\}} D_d^T \operatorname{Diag}(a_d) D_d \right)^{-1} \hat{l}. \quad (16)$$

2) *Solution to R-Problem (Reflectance Estimation):* After obtaining a desirable L , we compute R :

$$\begin{aligned} \operatorname{argmin}_R & \|R \circ L - S\|_F^2 + \beta \|\nabla R - G\|_F^2 \\ & + \omega \sum_i \|NN_i(R)\|_*. \end{aligned} \quad (17)$$

An alternating direction minimization algorithm [40] is provided to effectively optimize this objective function. First, by substituting R in the second term of Eqn. (17) with an auxiliary variable \hat{R} , the objective function can be rewritten as the following equivalent form:

$$\begin{aligned} \operatorname{argmin}_{R, \hat{R}} & \|R \circ L - S\|_F^2 + \omega \sum_i \|NN_i(R)\|_* \\ & + \beta \|\nabla \hat{R} - G\|_F^2, \quad \text{s. t. } R = \hat{R}. \end{aligned} \quad (18)$$

Then, a Lagrange multiplier is introduced to remove the equality constraint. The obtained augmented Lagrangian function of Eqn. (18) is as follows:

$$\begin{aligned} \mathcal{L}(R, \hat{R}, Z) = & \|R \circ L - S\|_F^2 + \omega \sum_i \|NN_i(R)\|_* \\ & + \beta \|\nabla \hat{R} - G\|_F^2 + Z \circ (\hat{R} - R) + \frac{\mu}{2} \|\hat{R} - R\|_F^2, \end{aligned} \quad (19)$$

where Z is the Lagrange multiplier and μ is a positive scalar.

Therefore, the original minimization problem (6) can be solved using standard alternating direction minimization solutions. Here we iteratively update each variable while regarding other variables that have been estimated in the previous iteration as constants.

2-1) Solution to $R\hat{R}$ Problem (Contrast Enhancement): Neglecting the terms unrelated to \hat{R} , we have the following optimization problem:

$$\begin{aligned} \operatorname{argmin}_{\hat{R}} & \beta \|\nabla \hat{R} - G\|_F^2 + Z^{(k)} \circ (\hat{R} - R^{(k)}) \\ & + \frac{\mu^{(k)}}{2} \|\hat{R} - R^{(k)}\|_F^2. \end{aligned} \quad (20)$$

Similar to the solution to Eqn. (14), the estimated \hat{R} of the $(k+1)$ -th iteration can be figured out by differentiating Eqn. (20) with respect to \hat{R} and setting the derivative to 0. Namely, we have:

$$\begin{aligned} & \left(2\beta \sum_{d \in \{h, v\}} D_d^T D_d + \mu I \right) \hat{r} \\ & = 2\beta \sum_{d \in \{h, v\}} D_d^T g_d + \mu^{(k)} r^{(k)} - z^{(k)}, \end{aligned} \quad (21)$$

where we use the small letter of r , g and z as the same meaning of the illumination estimation section. The estimated \hat{R} is:

$$\begin{aligned} \hat{r}^{(k+1)} &= \left(2\beta \sum_{d \in \{h, v\}} D_d^T D_d + \mu I \right)^{-1} \\ &\times \left(2\beta \sum_{d \in \{h, v\}} D_d^T g_d + \mu^{(k)} r^{(k)} - z^{(k)} \right), \end{aligned} \quad (22)$$

where \times means matrix multiplication.

2-2) Solution to $R\cdot R$ Problem (Noise Suppression): Collecting the terms related to R leads to the following problem:

$$\begin{aligned} \operatorname{argmin}_R & \|R \circ L - S\|_F^2 + \omega \sum_i \|NN_i(R)\|_* \\ & + Z^{(k)} \circ (\hat{R}^{(k+1)} - R) + \frac{\mu^{(k)}}{2} \|\hat{R}^{(k+1)} - R\|_F^2. \end{aligned} \quad (23)$$

Note that $\{NN_i(R)\}$ are corresponded to R but have a different form: R is an image-level representation and $\{NN_i(R)\}$ are patch-based ones. For simplicity, we reform Eqn. (23) into:

$$\begin{aligned} \operatorname{argmin}_R & \sum_i \|NN_i(R) \circ L_i - S_i\|_F^2 + \omega \sum_i \|NN_i(R)\|_* \\ & + \sum_i Z_i^{(k)} \circ (\hat{R}_i^{(k+1)} - NN_i(R)) \\ & + \sum_i \frac{\mu^{(k)}}{2} \|\hat{R}_i^{(k+1)} - NN_i(R)\|_F^2, \end{aligned} \quad (24)$$

where L_i , S_i , $Z_i^{(k)}$, $\hat{R}_i^{(k+1)}$ are the patch-level representations of L , S , $Z^{(k)}$, $\hat{R}^{(k+1)}$, respectively.

To make it more concise, we now consider each location and omit all operators with respect to i in the following presentation. Now we have:

$$\begin{aligned} \operatorname{argmin}_R & \|R \circ L - S\|_F^2 + \omega \|R\|_* \\ & + Z^{(k)} \circ (\hat{R}^{(k+1)} - R) + \frac{\mu^{(k)}}{2} \|\hat{R}^{(k+1)} - R\|_F^2. \end{aligned} \quad (25)$$

We first simplify the rest part of the optimization function apart from the nuclear norm. The above expression can be reformed as:

$$\begin{aligned} \operatorname{argmin}_R & \|R\|_* + \frac{L^2}{\omega} \circ \|R - S/L\|_F^2 \\ & + \frac{\mu^{(k)}}{2\omega} \|R - (\hat{R}^{(k+1)} + Z^{(k)}/\mu^{(k)})\|_F^2. \end{aligned} \quad (26)$$

It is a modified low-rank minimization problem and can be transformed into the following formation:

$$\begin{aligned} \operatorname{argmin}_R & \|R - \bar{R}^{(k)}\|_F^2 + \|R\|_*, \\ \bar{R}^{(k)} &= \frac{2S \circ L + \mu^{(k)} \hat{R}^{(k+1)} + Z^{(k)}}{2L^2 + \mu^{(k)}}. \end{aligned} \quad (27)$$

Therefore, the original $R\cdot R$ problem in Eqn. (23) now turns into a standard low-rank minimization problem [41]:

$$\begin{aligned} \operatorname{argmin}_{NN_i(R)} & \|NN_i(R) - NN_i(\bar{R}^{(k)})\|_F^2 + \|NN_i(R)\|_*^2, \\ \bar{R}^{(k)} &= \frac{2S \circ L + \mu^{(k)} \hat{R}^{(k+1)} + Z^{(k)}}{2L^2 + \mu^{(k)}}, \end{aligned} \quad (28)$$

which has a closed-form solution:

$$NN_i(\bar{R}^{(k+1)}) = S_\tau(\bar{R}^{(k)}), \quad (29)$$

where S_τ is the soft shrinkage process.

2-3) Updating Z and μ : After that, the auxiliary matrix Z and the penalty scalar μ are updated through:

$$Z^{(k+1)} = Z^{(k)} + \mu^{(k)} (\hat{R}^{(k+1)} - R^{(k+1)}). \quad (30)$$

$$\mu^{(k+1)} = \mu^{(k)} \rho, \quad \rho > 1. \quad (31)$$

Algorithm 1 Robust Single-Image Low-Light Enhancement

Input: The input image S .

- 1: Estimate \hat{L} via Eqn. (11);
– Illumination Initialization
- 2: Estimate L via Eqn. (16);
– Illumination Estimation
- 3: $R^{(0)} = S/L$, $Z^{(0)} = 0$, $\mu^{(0)} = 1$, $\rho = 1.5$, $k = 0$;
– Reflectance Initialization
- 4: **while** not converged **do**
- 5: Update $\hat{R}^{(k+1)}$ via Eqn. (22);
– Contrast Enhancement
- 6: Update $R^{(k+1)}$ via Eqn. (29);
– Noise Suppression
- 7: Update $Z^{(k+1)}$ via Eqn. (30);
- 8: Update $\mu^{(k+1)}$ via Eqn. (31).
- 9: **end while**
- 10: Estimate S' via Eqn. (32).
– Illumination Adjustment

Output: The estimated result S' .

3) Illumination Adjustment: After estimating the illumination L and reflectance R , the gamma correction is applied to

adjust the illumination. Then, the final enhancement result S' is generated by:

$$S' = R \circ L'^{\frac{1}{\gamma}}, \quad (32)$$

where L' is the normalized L , and γ is empirically set to 2.2.

With an input noisy low-light image, the whole process of our method for image enhancement and denoising is summarized in Algorithm 1.

C. Video Low-Light Enhancement

In this part, we extend our LR3M model to video low-light enhancement. Compared to single image low-light enhancement, video enhancement has two additional key elements:

- Temporal coherence. The extracted features, *e.g.* illumination and reflectance, should be continuous along the motion trajectories among frames.
- Temporal redundancy. In video enhancement, there are abundant temporal correlation and structural correspondence to facilitate inferring information lost in the degradation process.

Based on these considerations, besides combining the optimization function of each frame, the optimization function of LR3M for video enhancement makes the following changes. First, a coherence term is introduced to enforce the illumination consistency among frames. Second, in the solution of R - R problem, the similar patches are also searched from the adjacent frames. The overall function turns to:

$$\begin{aligned} \operatorname{argmin}_{\mathbb{L}, \mathbb{R}} & \sum_{t=1}^n \sum_{j=1}^{t-1} \|L_t - F_{t,j}(L_j)\|_F^2 + \sum_{t=1}^n \left(\|R_t \circ L_t - S_t\|_F^2 \right. \\ & \left. + \alpha \|\nabla L_t\|_1 + \beta \|\nabla R_t - G_t\|_F^2 + \omega \sum_i \|NN_i(R_t)\|_* \right), \end{aligned} \quad (33)$$

where $\mathbb{S} = \{S_1, S_2, \dots, S_n\}$, $\mathbb{L} = \{L_1, L_2, \dots, L_n\}$ and $\mathbb{R} = \{R_1, R_2, \dots, R_n\}$ denote sequence of input frames, the estimated illumination maps and reflectance maps, respectively. n represents the number of total input frames and function $C(\cdot)$ measures the inconsistency of input sequence. $F_{t,j}(\cdot)$ builds the location correspondence between L_t and L_j .

Note that when dealing with a single image, n equals to 1. There is no inconsistency in the input sequence, which means the value of function $C(L)$ is 0. The objective function in Eqn. (33) is simplified to Eqn. (6).

To solve the problem in Eqn. (33), we slightly change the formation of *L*-problem (illumination estimation) as follows:

$$\begin{aligned} \operatorname{argmin}_{\mathbb{L}, \mathbb{R}} & \sum_{t=1}^n \sum_{j=1}^{t-1} \|L_t - F_{t,j}(L_j)\|_F^2 + \\ & \|R_t \circ L_t - S_t\|_F^2 + \alpha \|\nabla L_t\|_1. \end{aligned} \quad (34)$$

We use a very simple method to solve this problem: first estimating $\{L_t\}$ based on Eqns. (11) and (16), then smoothing $\{L_t\}$ to obtain $\{I_t\}$ based a joint spatial and temporal smoothing with structure preservation. We use the guided filter [42] to smooth every 2D slice of the illumination of successive

Algorithm 2 Robust Video Low-Light Enhancement

Input: The input frames $\mathbb{S} = \{S_1, S_2, \dots, S_n\}$.

- 1: Estimate $\mathbb{L} = \{L_1, L_2, \dots, L_n\}$ via Eqns. (11) and (16);
- 2: Estimate $\mathbb{I} = \{I_1, I_2, \dots, I_n\}$ via Eqn. (35);
 - *Illumination Smoothing*
- 3: Estimate $\mathbb{R} = \{R_1, R_2, \dots, R_n\}$ via Eqns. (22) – (31);
 - *Consider \mathbb{L} as \mathbb{I}*
- 4: Estimate $\mathbb{S}' = \{S'_1, S'_2, \dots, S'_n\}$ via Eqn. (32).

Output: The estimated result \mathbb{S}' .



Fig. 3: Dataset used in this paper, which contains 24 low-light images with different noise levels.

frames under the guidance of input sequence iteratively, which is equivalent to 3D structural preserving smoothing. This procedure is described as follows:

$$\mathbb{I} = F(\mathbb{L}; \mathbb{S}), \quad (35)$$

where $\mathbb{I} = \{I_1, I_2, \dots, I_n\}$ denotes the sequence of smoothed illumination maps. $F(\cdot)$ represents the filtering function with \mathbb{L} as input and \mathbb{S} as guidance.

In the reflectance estimation, we use smoothed I to replace estimated L in Eqn. (17) and adjust the block matching function so that it can find nearest blocks through adjacent frames. The whole process of our LR3M for video low-light enhancement and denoising is summarized in Algorithm 2.

IV. EXPERIMENTAL RESULTS

In this section, the experimental results and our analyses are presented. We first give some analyses of our method. Then, the proposed method is compared with existing state-of-the-art methods quantitatively and qualitatively. Furthermore, we exhibit the results of our method in video enhancement tasks to demonstrate the generality of our model.

A. Experimental Settings

All experiments are run on MATLAB R2014a with 4G RAM and Intel Core i5-4210H CPU @2.90GHz. The parameters α and β in Eqn. (6) are empirically set as 0.015 and 2. ϵ and λ are set to be 1 and 3.5. In general cases, this setting performs well.

We test 24 low-light images as shown in Fig. 3. This dataset contains images provided by authors of [16, 19] and some challenging data with much noise from [43] and [44].

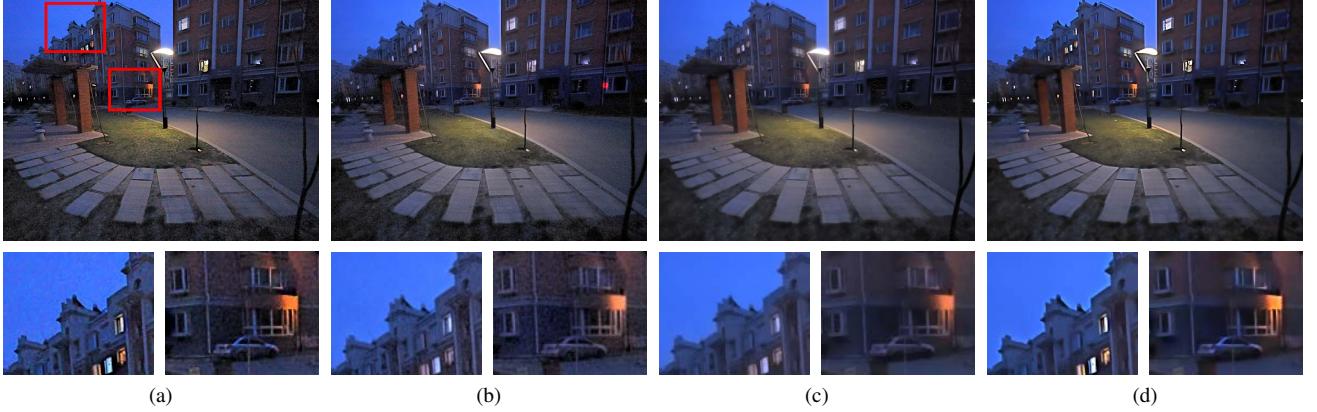


Fig. 4: Comparison of different settings. (a) is the result without low-rank term. (b) is the result with $\alpha = 0$ and $\beta = 0$. The result with $\beta = 0$ is (c). The result of our method is (d).

The dataset of [16] contains 10 low-light images. Most of its images are noise-free. The dataset of [19] contains 18 low-light images, which includes both landscape and human images. DICM [43] dataset contains 69 captured images from commercial digital cameras. Among them, 44 images suffer from low illumination. Authors of [44] provided some noisy videos and some of them were obtained in low-light condition. We use their data for video enhancement and use extracted frames for image enhancement. In summary, our dataset is a collection of indoor/outdoor low-light images with different noise levels.

B. Analyses of Our Method

1) *Identification of the unified framework:* In the first place, we give some analyses of our method. We argue that in terms of noisy low-light images, it is better to perform enhancing and denoising simultaneously than to separate them into two independent jobs. Because it is problematic to decide the order of these two jobs. To demonstrate that, we use LIME [16] and BM3D [17] to solve the same task by enhancing and denoising the low light images separately. In addition, the parameter σ used in BM3D [17] which indicates the noise level is set as 5 if the input is denoised first and 10 otherwise. As is shown in Fig. 2, if denoising comes before enhancing, the denoising procedure easily blurs the image and causes inaccuracy. On the other hand, if the order is reversed, the noise in low-light image is inevitably amplified so that the effectiveness of the following denoising process is weakened, although σ is much larger.

2) *Sequential decomposition:* To show the effectiveness of sequential decomposition, we use two methods to solve our objective function, *i.e.*, the iterative solution and sequential solution. For iterative solution, we substitute ∇L in the second term of Eqn. (6) with an auxiliary variable T . Then we have a corresponding augmented Lagrangian function:

$$\begin{aligned} \mathcal{L}(L, R, T, Y) = & \|R \circ L - S\|_F^2 + \alpha \|T\|_1 \\ & + Y \circ (\nabla L - T) + \frac{\eta}{2} \|\nabla L - T\|_F^2 \\ & + \omega \sum_i \|NN_i(R)\|_* + \beta \|\nabla R - G\|_F^2, \quad (36) \end{aligned}$$

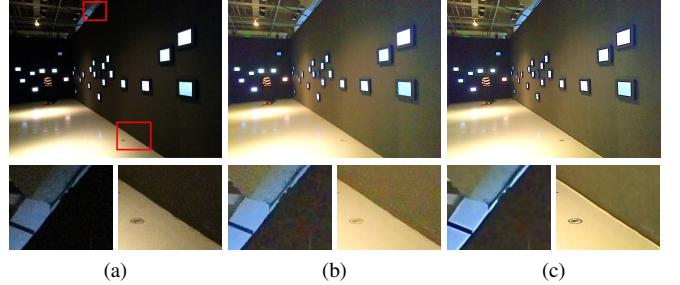


Fig. 6: Comparison of an iterative decomposition procedure (b) and a sequential one (c). From left to right: the input image, result of the iterative solution to Eqn. (6) and result of our method (LR3M).

where Y is the Lagrange multiplier and η is a positive scalar.

In every iteration, we estimate $L^{(k+1)}$ by:

$$\begin{aligned} l^{(k+1)} = & \left(2r^{(k)}{}^T r^{(k)} + \eta \sum_{d \in \{h, v\}} D_d^T D_d \right)^{-1} \\ & \times \left((2r^{(k)} s + \eta D^T(t^{(k)} - \frac{y^{(k)}}{\eta})) \right). \end{aligned} \quad (37)$$

Then we consider $L^{(k+1)}$ as constant and estimate $R^{(k+1)}$ as we discussed in Sec III. After that we update T , y and η using:

$$T^{(k+1)} = S_{\frac{\alpha}{\eta^{(k)}}}(\nabla L^{(k+1)} + \frac{Y^{(k)}}{\eta^{(k)}}). \quad (38)$$

$$Y^{(k+1)} = Y^{(k)} + \eta^{(k)}(\nabla L^{(k+1)} - T^{(k+1)}). \quad (39)$$

$$\eta^{(k+1)} = \eta^{(k)}\rho, \quad \rho > 1. \quad (40)$$

The results using two different solutions are shown in Fig. 6. We can see that the sequential solution obtains a better result than the iterative one, mainly because it prevents the illumination map from being impaired and makes sure that most noise is left in the reflectance map.

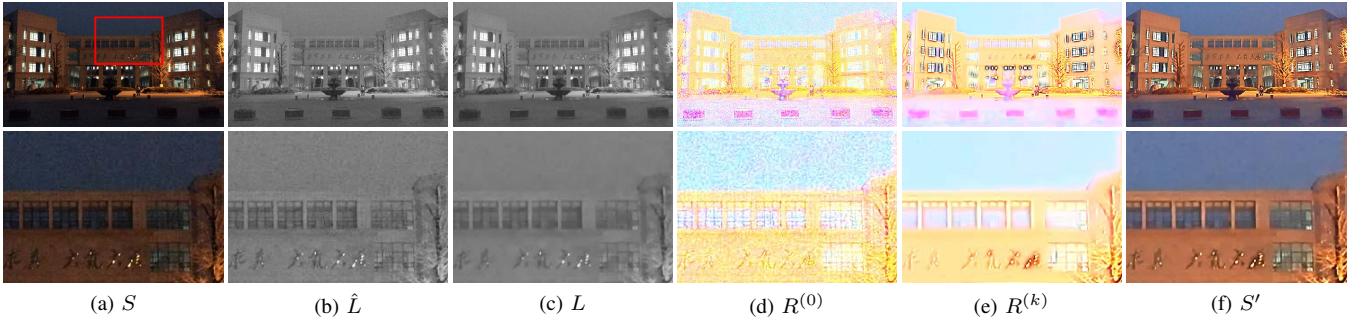


Fig. 5: Decomposition results and corresponding enhancement results of our method. From left to right: the original image S , the initial illumination map \hat{L} and estimated illumination map L , a rough reflectance map computed by $R^{(0)} = S/L$ using element-wise division, the estimated reflectance map $R^{(k)}$ after k iterations and the final result S' . For better visual quality, \hat{L} and L are transformed by gamma transformation.

3) *Parameter study*: We also study the effectiveness of different terms in our objective function. The results of different objective function settings are displayed in Fig. 4. Group (a) shows the result without the low-rank term. Since low-rank minimization is the key to remove noise while enhancing the whole image, missing this term leads to a noisy result. Much noise can be observed in the enlarged local areas shown below the result image. The second column (b) is the result with $\alpha = 0$, which means the second term of Eqn. (6) is inactive. Without this term restricting total variation sparsity, the illumination component becomes more noisy and breaks the assumption of illumination smoothness. It results in a worse result compared to (d), the proposed method with α greater than zero. Besides, removing this term also causes some mistakes in the output image, *i.e.* the red part in the right building, in the first row of (b). (c) is obtained with $\beta = 0$. It is much clearer and free of noise compared to the former ones. However, without the effort to amplify the contrast of the image, it looks less distinct and visually pleasing compared to our proposed method (d), which not only removes noise when enhancing the image, but also improves the contrast of it. Consequently, all the constraints in Eqn. (6) are necessary for deriving a noise free and visually pleasing result.

4) *Low-rank regularized Retinex model*: Fig. 5 shows an example of the decomposition results using our method. From (b) and (c) we can see that after solving Eqn. (9), we obtain a much more preferable illumination map. It means that when we amplify the estimated illumination, noise is not enhanced. Through (c) and (d) we note that with an original input image full of noise, after we obtain a clean illumination map, most noise is left in the other part of Retinex model, *i.e.* the initial reflectance map shown in (d). This phenomenon agrees with our assumption. Compared to (d), (e) is less noisy thanks to the low-rank restriction in our method. Furthermore, we also improve the overall contrast. This can be observed at the edges of windows and the tree in (e). Collecting all these refinements, our final result is demonstrated in (f). It is clearer and lighter than the input image. In total, our method brings three benefits, *i.e.*, enhancing the illumination of a low-light image, preventing noise amplifying and further removing it, and improving the contrast and visual quality of the final result.

C. Qualitative Comparison

1) *Low-light Enhancement Effectiveness*: To evaluate the enhancing effectiveness of our proposed method, we compare it with conventional histogram equalization (HE), contrast-limited adaptive histogram equalization (CLAHE) [3], bio-inspired multi-exposure fusion framework (BIMEF) [24] and state-of-the-art Retinex based enhancement methods, *i.e.*, simultaneous reflectance and illumination estimation (SRIE) [15], naturalness preserved enhancement algorithm (NPE) [18], probabilistic method for image enhancement (PIE) [39].

HE and CLAHE are performed using the MATLAB built-in functions *histeq* and *adapthisteq*. They stretch the narrowly distributed histograms of low-light images globally or locally in order to enhance the contrast. The results of other methods are generated by the code downloaded from the authors websites, with recommended experiment settings. BIMEF fuses multi-exposure images using illumination estimation techniques to provide an accurate contrast and lightness enhancement. SRIE uses a weighted variational model to avoid disadvantages of logarithmic transformation. NPE is designed to preserve the naturalness of images, and most of its results have vivid color. PIE uses probabilistic method to simultaneous estimation of illumination and reflectance in the linear domain.

Figs. 7 and 8 show several comparisons between enhancement results generated by different methods. In Fig. 7, HE, CLAHE, SRIE, PIE and BIMEF all fail to enhance the right window and NPE amplifies noise when enhancing it. Besides, the result of BIMEF suffers from color distortion. The center of the girl's dress is not enhanced in (c), with grey areas even darker than the input image. On the other hand, our method not only enhances this part, but also keeps its difference from left part of the dress, unlike HE, which makes the entire dress under the same illumination level. In the second rows of (b) and (e), the girl's dress loses texture details in the left part. Compared to them, our method can generate a pleasing result.

In Fig. 8, similarly, HE, SRIE, NPE PIE and BIMEF erase details in the wall of the tower and the house while CLAHE makes the wall darker. Looking at the sky in Fig. 8, it is obvious that methods like HE, CLAHE and BIMEF cannot



Fig. 7: Comparisons of low-light image enhancement results.

handle it correctly. What's more, compared to (d) and (f), our result is enhanced more successfully, since in the second row, the house of ours is of better visibility than those of SRIE, PIE and BIMEF.

2) Denoising Effectiveness: Besides enhancing effectiveness, our method also demonstrates an impressive denoising ability. Figs 9, 10 show comparisons among our methods and two enhancing methods considering noise. Namely, LIME [16] and structure-revealing low-light image enhancement [19].

Results of LIME are generated following their work with a post denoising procedure of BM3D [17], as mentioned in their paper. In general, LIME does not handle the problem of illumination over-enhancement, which adds to the noise

amplification. For example, in the second column of Fig. 9, it can be observed that the car is over-enhanced and in (b) of Fig. 10, the light from the windows is too bright and the light on the desk exceeds its original region. It makes noise obvious despite using an advanced denoising method.

The method of Li *et al.* [19] particularly takes noise into consideration under a unified framework rather than similarly removing the noise by post processing. But its results still have some noise left and sometimes even more noise than results of LIME, as can be seen in (c) of Fig. 9, there is more noise around the light. This is because they used an iterative method to solve their objective function. A weighted matrix is introduced to restrict noise in the reflectance map. However,



Fig. 8: Comparisons of low-light image enhancement results.

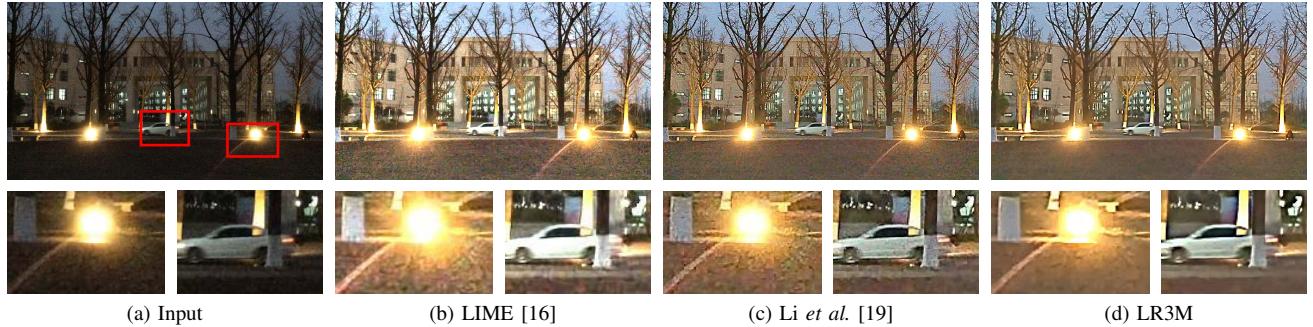


Fig. 9: Comparisons of low-light image enhancement results with a noisy input.

during the iterations, the noise escapes from the reflectance to the illumination. Furthermore, in some cases where there is much noise, the restricted matrix is not enough to handle it. In contrast, we use the low-rank regulation to effectively remove noise.

Fig. 11 gives a more closed comparison on denoising effectiveness. Since results of LIME are always noisier than those of the other two methods, we only compare our proposed method with the method of [19] in Fig. 11. We can find that our approach performs better in all cases shown in Fig. 11, which gives a convincing proof of the superiority of our method.

D. Quantitative Comparison

In addition to subjective comparisons, we also apply objective measurements to evaluate the performance of the proposed method objectively. Considering the fact that low-light images seldom have corresponding reference images, we adopt three

blind quality assessments. Namely, blind tone-mapped quality index (BTMqi) [45], no-reference image quality metric for contrast distortion (NIQMC) [46] and blind sharpness metric in the autoregressive parameter space (ARISM) [47] to

TABLE I: Objective Evaluation on Our Dataset

Method	BTMQI↓	ARISM C ↓	NIQMC↑	Ave. Rank
Input	6.3792	2.9037	4.1481	10
HE	3.9203	3.1459	5.4059	5
CLAHE [3]	4.8053	2.9050	4.9644	6
SRIE [15]	4.5550	2.9483	4.7257	9
NPE [18]	3.7687	3.0376	4.8105	6
PIE [39]	4.6902	2.9024	4.7989	8
BIMEF [24]	3.6636	2.9689	4.9413	4
LIME [16]	3.8150	2.8319	5.4161	2
Li <i>et al.</i> [19]	3.7715	2.8621	4.9256	3
LR3M	3.7563	2.8305	5.0071	1

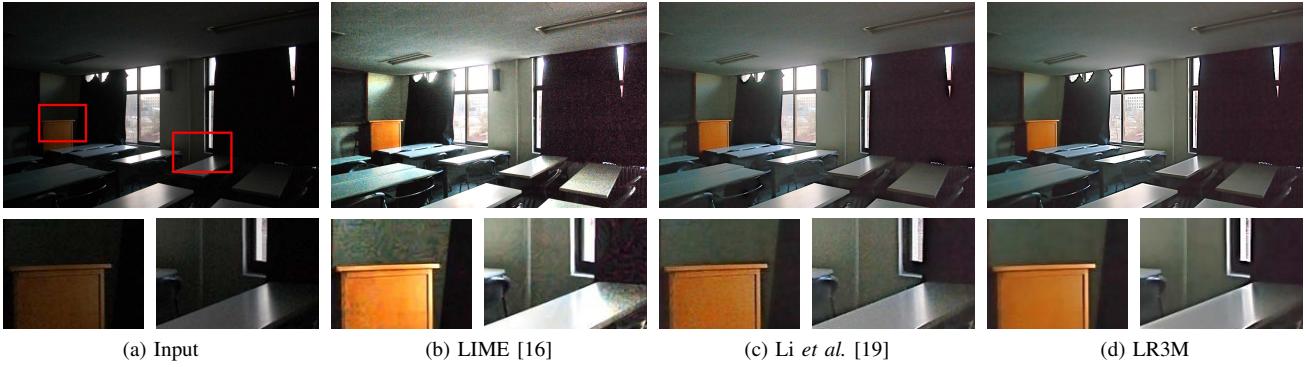


Fig. 10: Comparisons of low-light image enhancement results with a noisy input.

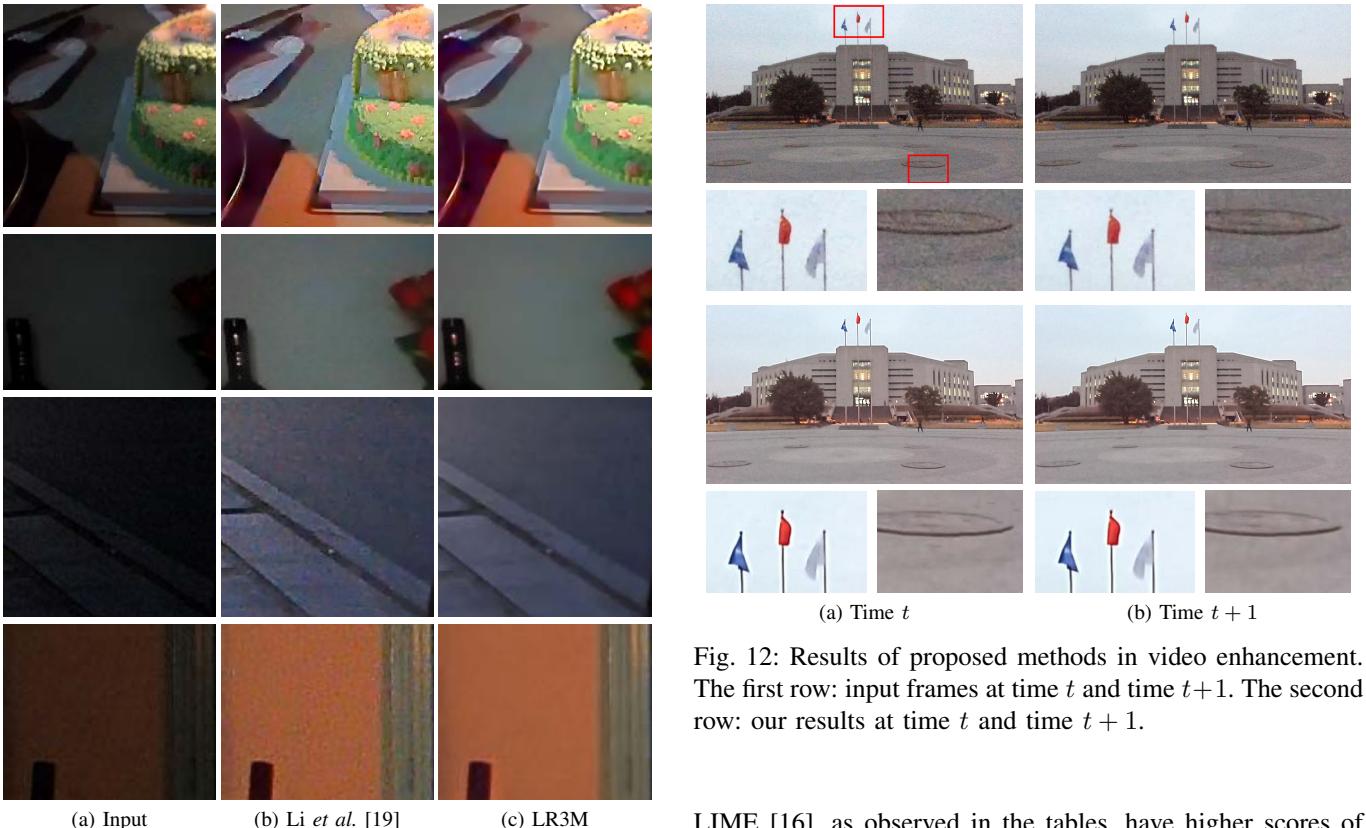
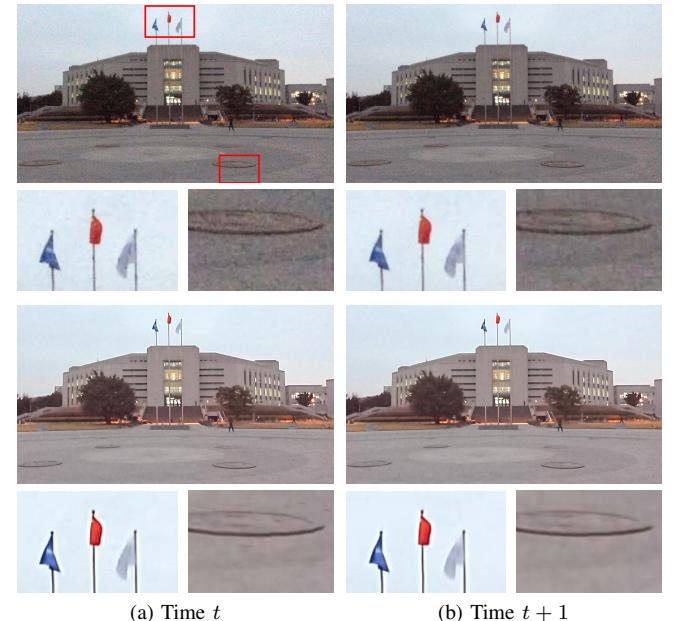


Fig. 11: Comparisons of low-light image enhancement results in detail.

evaluate the enhancement results comprehensively. Results of different approaches are summarized in Tab. I.

We assess objective qualities on datasets of [16] and [19]. For BTMQI and ARISMC, smaller values represent better image qualities and for NIQMC, larger values indicate better qualities of image contrast. BTMQI measures image quality by measuring the average intensity, contrast, and structure information of tone-mapped images. ARISMC estimates the image sharpness considering both luminance and chromatic components, *a.k.a.*, illumination and reflectance. As for NIQMC, it favors images with higher contrast.

From Tab. I we can see that results of our method generally have the best qualities on both datasets. Methods of HE and

Fig. 12: Results of proposed methods in video enhancement. The first row: input frames at time t and time $t+1$. The second row: our results at time t and time $t+1$.

LIME [16], as observed in the tables, have higher scores of NIQMC, which indicates strong contrast, it is because they usually over-enhance input images. For example, the sky and wall in Fig. 8 (b) and the light and car in Fig. 9 (b). The results of our methods rank within the top three for all objective measurements. Therefore, our method has the top average rank, estimated by adding up the ranks of every measurements.

E. Results of Video Enhancement

Our robust low-light enhancement method can also be applied in video tasks and show impressive results. Fig. 12 shows our enhancing results of two successive frames denoted as the frame at time t and $t+1$. Our method shows impressive denoising effectiveness and the obtained results are temporally consistent. Moreover, the illumination of the original video is not very low. Our method is able to adaptively decide the enhance level and avoid over-enhancement.



Fig. 13: Two visual comparisons of sequential and iterative estimations.

F. Visual Comparison of Sequential and Iterative Estimations

As discussed in [19], for iterative methods, the fidelity term might play a misleading role in the whole estimation. During the iteration process, intensive noise hidden in the observed image S will be assigned to either R or L . In most previous methods, efforts are put into the illumination component L , and the reflectance is estimated by $R = S/L$, which inevitably includes noise and leads to a noisy result. We show two visual comparison results in Fig. 13. In the top panel, the input image is almost noise-free, while in the bottom panel, the input image contains intensive noise. It is observed that, for the noisy input image, the sequential method achieves better visual quality. Comparatively, the iterative method leads to a noisy enhanced result.

G. Comparison to Deep Prior Ensemble

We compare the visual results of our method with Deep Prior Ensemble (DPE) [48]. As shown in Fig. 14, the results of DPE have good visual quality. Comparatively, our methods are better in noise suppression and contrast stretching.

H. Limitations of Low-Rank Prior

First, it brings in more computation complexity to adopt the low-rank prior, because it requires to calculate the singular value decomposition of matrices. Second, since the low-rank minimization is a patch-based algorithm, it will be less effective if these are no similar patches and the noise level is low. We show one failure case in the region full of textures as shown in Fig. 15. It is observed that, with the low-rank prior, the structural details are enhanced, *e.g.* the contours between the grass and flowers and those of sands. However, some internal textures are removed, since in these regions,



Fig. 14: Visual comparisons of DPE [48] and the proposed method. Zooming-in the figure will provide a better look at the enhancement quality.



Fig. 15: Visual comparisons of the results with and without low-rank priors.

the similar patches are hard to be obtained. It is interesting in the future to dynamically decide the weighting of the low-rank term in the joint optimization based on the regional noise level and patch-level redundancy.

I. Ablation Studies on Low-Rank Prior

Beyond the traditional two-dimensional (2D) low rank prior to suppress noise, we propose a three-dimensional (3D) low rank prior in the video low-light enhancement case, where each atom signal is flattened from a pixel volume $R_{it} \in \mathbb{R}^{b^2 \times s}$ considering the temporal context to better stabilize the estimation of the illumination and reflectance layers along the temporal dimension. b and s are the spatial patch size, and temporal length. t indexes the t -th most similar atom in the search window. This 3D low rank prior leads to more temporally consistent results. We show an example of the visual results without the low-rank prior, with the traditional 2D low-rank prior, and with our proposed 3D low-rank prior in Fig. 16. It is observed that, our method is superior in both noise suppression and temporal consistency. Especially for the regions denoted

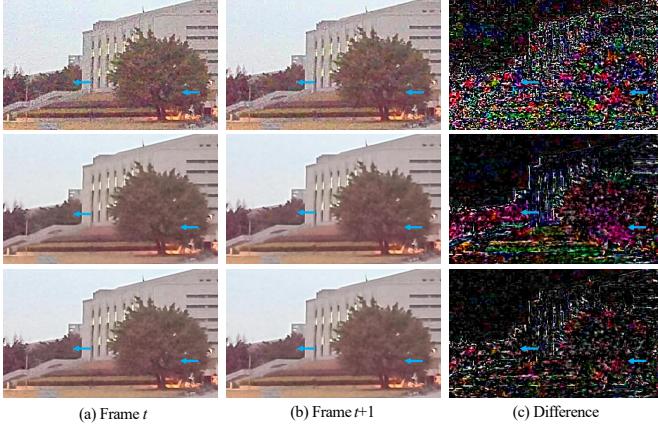


Fig. 16: An example of visual results without the low-rank prior (top panel), with the traditional 2D low-rank prior (middle panel), and with our proposed 3D low-rank prior (bottom panel).

by blue arrows, it is observed that, our proposed 3D low-rank prior in the bottom panel leads to the sparsest and smallest temporal differences, which demonstrates our superiority in the temporal consistency.

J. Comparison to Learning-Based Denoising Methods

We compare our method with learning-based denoising methods, including DnCNN [49] and FFDNet [50], as shown in Fig. 17. The denoising operations are applied on the reflectance layers. In general, the results include remaining noise (denoted by red arrows) and blurred details (denoted by blue arrows), as these methods are trained on nature images rather than the reflectance images. After extracting the illumination, the noise in the reflectance is amplified and has a different distribution from that in the original image.

V. CONCLUSION

In this paper we discuss the existing problem of noise in mainstream methods of low-light enhancement domain. And we argue that noise should be particularly handled during low-light enhancement. However, existing methods either ignore this issue or do not handle it well. According to that, we constructively present a robust low-light enhancement method via low-rank regularized Retinex model. By intentionally limiting noise to the minimum, we obtain high-quality images finally. Our experiments show the impracticality of treating denoising as a pre/post processing procedure and extensive results demonstrate the effectiveness of our method.

REFERENCES

-
- [1] C. Chen, Q. Chen, J. Xu, and V. Koltun, “Learning to see in the dark,” in *IEEE Conference on Computer Vision and Pattern Recognition*, June 2018.
- [2] M. Abdullah-Al-Wadud, M. H. Kabir, M. A. A. Dewan, and O. Chae, “A dynamic histogram equalization for image contrast enhancement,” *IEEE Transactions on Consumer Electronics*, vol. 52, no. 2, pp. 593–600, May 2007.
- [3] S. M. Pizer, R. E. Johnston, J. P. Erickson, B. C. Yankaskas, and K. E. Muller, “Contrast-limited adaptive histogram equalization: speed and effectiveness,” *Proceedings of the First Conference on Visualization in Biomedical Computing*, pp. 337–345, May 1990.
- [4] L. Li, R. Wang, W. Wang, and W. Gao, “A low-light image enhancement method for both denoising and contrast enlarging,” *IEEE International Conference on Image Processing*, pp. 3730–3734, September 2015.
- [5] L. Zhang, P. Shen, X. Peng, G. Zhu, J. Song, W. Wei, and H. Song, “Simultaneous enhancement and noise reduction of a single low-light image,” *IET Image Processing*, vol. 10, no. 11, pp. 840–847, 2016.
- [6] D. J. Jobson, Z. Rahman, and G. A. Woodell, “Properties and performance of a center/surround retinex,” *IEEE Transactions on Image Processing*, vol. 6, no. 3, pp. 451–462, March 1997.
- [7] ——, “A multiscale retinex for bridging the gap between color images and the human observation of scenes,” *IEEE Transactions on Image Processing*, vol. 6, no. 7, pp. 965–976, July 1997.
- [8] M. Herscovitz and O. Yadid-Pecht, “A modified multi scale retinex algorithm with an improved global impression of brightness for wide dynamic range pictures,” *Machine Vision and Applications*, vol. 15, no. 4, pp. 220–228, October 2004.
- [9] C. Xiao and Z. Shi, “Adaptive bilateral filtering and its application in retinex image enhancement,” *Seventh International Conference on Image and Graphics*, pp. 45–49, July 2013.
- [10] J. Yang, X. Jiang, C. Pan, and C.-L. Liu, “Enhancement of low light level images with coupled dictionary learning,” *International Conference on Pattern Recognition*, pp. 751–756, December 2016.
- [11] K. G. Lore, A. Akintayo, and S. Sarkar, “Llnet: A deep autoencoder approach to natural low-light image enhancement,” *Pattern Recognition*, vol. 61, pp. 650 – 662, 2017.
- [12] R. Kimmel, M. Elad, D. Shaked, R. Keshet, and I. Sobel, “A variational

- framework for retinex,” *International Journal of Computer Vision*, vol. 52, no. 1, pp. 7–23, April 2003.
- [13] M. Ng and W. Wang, “A total variation model for retinex,” *SIAM Journal on Imaging Sciences*, vol. 4, no. 1, pp. 345–365, 2011.
- [14] X. Fu, Y. Sun, M. LiWang, Y. Huang, X. Zhang, and X. Ding, “A novel retinex based approach for image enhancement with illumination adjustment,” *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 1190–1194, May 2014.
- [15] X. Fu, D. Zeng, Y. Huang, X. P. Zhang, and X. Ding, “A weighted variational model for simultaneous reflectance and illumination estimation,” *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2782–2790, June 2016.
- [16] X. Guo, Y. Li, and H. Ling, “LIME: Low-light image enhancement via illumination map estimation,” *IEEE Transactions on Image Processing*, vol. 26, no. 2, pp. 982–993, February 2017.
- [17] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, “Image denoising by sparse 3-D transform-domain collaborative filtering,” *IEEE Transactions on Image Processing*, vol. 16, no. 8, pp. 2080–2095, August 2007.
- [18] S. Wang, J. Zheng, H. M. Hu, and B. Li, “Naturalness preserved enhancement algorithm for non-uniform illumination images,” *IEEE Transactions on Image Processing*, vol. 22, no. 9, pp. 3538–3548, September 2013.
- [19] M. Li, J. Liu, W. Yang, X. Sun, and Z. Guo, “Structure-revealing low-light image enhancement via robust retinex model,” *IEEE Transactions on Image Processing*, vol. 27, no. 6, pp. 2828–2841, June 2018.
- [20] X. Ren, M. Li, W. Cheng, and J. Liu, “Joint enhancement and denoising method via sequential decomposition,” *IEEE International Symposium on Circuits and Systems*, pp. 1–5, May 2018.
- [21] Z. Farbman, R. Fattal, D. Lischinski, and R. Szeliski, “Edge-preserving decompositions for multi-scale tone and detail manipulation,” *ACM Transactions on Graphics*, vol. 27, no. 3, pp. 67:1–67:10, August 2008.
- [22] S. Paris, S. W. Hasinoff, and J. Kautz, “Local laplacian filters: Edge-aware image processing with a laplacian pyramid,” *Communications of the ACM*, vol. 58, no. 3, pp. 81–91, February 2015.
- [23] M. Li, X. Wu, J. Liu, and Z. Guo, “Restoration of unevenly illuminated images,” *IEEE International Conference on Image Processing*, pp. 1118–1122, October 2018.
- [24] Z. Ying, G. Li, and W. Gao, “A bio-inspired multi-exposure fusion framework for low-light image enhancement,” *ArXiv: 1711.00591*, November 2017.
- [25] L. Shen, Z. Yue, F. Feng, Q. Chen, S. Liu, and J. Ma, “MSR-net:Low-light Image Enhancement Using Deep Convolutional Network,” *ArXiv: 1711.02488*, November 2017.
- [26] C. Wei, W. Wang, W. Yang, and J. Liu, “Deep retinex decomposition for low-light enhancement,” *British Machine Vision Conference*, September 2018.
- [27] W. Yang, S. Wang, Y. Fang, Y. Wang, and J. Liu, “From fidelity to perceptual quality: A semi-supervised approach for low-light image enhancement,” in *IEEE Conference on Computer Vision and Pattern Recognition*, June 2020.
- [28] E. J. Candès and B. Recht, “Exact matrix completion via convex optimization,” *Foundations of Computational Mathematics*, vol. 9, no. 6, pp. 717–772, December 2009.
- [29] J. Wright, A. Ganesh, S. Rao, Y. Peng, and Y. Ma, “Robust principal component analysis: Exact recovery of corrupted low-rank matrices via convex optimization,” *Advances in Neural Information Processing Systems* 22, pp. 2080–2088, 2009.
- [30] Z. Lin, A. Ganesh, J. Wright, L. Wu, M. Chen, and Y. Ma, “Fast convex optimization algorithms for exact recovery of a corrupted low-rank matrix,” *Intl. Workshop on Comp. Adv. in Multi-Sensor Adapt. Processing*, 2009.
- [31] D. Zhang, Y. Hu, J. Ye, X. Li, and X. He, “Matrix completion by truncated nuclear norm regularization,” *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2192–2199, June 2012.
- [32] Y. Hu, D. Zhang, J. Ye, X. Li, and X. He, “Fast and accurate matrix completion via truncated nuclear norm regularization,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 9, pp. 2117–2130, September 2013.
- [33] S. Ono, T. Miyata, and I. Yamada, “Cartoon-texture image decomposition using blockwise low-rank texture characterization,” *IEEE Transactions on Image Processing*, vol. 23, no. 3, pp. 1128–1142, March 2014.
- [34] W. Dong, G. Shi, and X. Li, “Nonlocal image restoration with bilateral variance estimation: A low-rank approach,” *IEEE Transactions on Image Processing*, vol. 22, no. 2, pp. 700–711, February 2013.
- [35] M. Li, J. Liu, Z. Xiong, X. Sun, and Z. Guo, “Marlow: A joint multiplanar autoregressive and low-rank approach for image completion,” *IEEE European Conference on Computer Vision*, 2016.
- [36] M. Li, J. Liu, X. Sun, and Z. Xiong, “Image/video restoration via multiplanar autoregressive model and low-rank optimization,” *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 15, no. 4, Dec. 2019.
- [37] M. Ng and W. Wang, “A total variation model for retinex,” *SIAM Journal on Imaging Sciences*, vol. 4, no. 1, pp. 345–365, 2011.
- [38] L. Wang, L. Xiao, H. Liu, and Z. Wei, “Variational bayesian method for retinex,” *IEEE Transactions on Image Processing*, vol. 23, no. 8, pp. 3381–3396, August 2014.
- [39] X. Fu, Y. Liao, D. Zeng, Y. Huang, X. P. Zhang, and X. Ding, “A probabilistic method for image enhancement with simultaneous illumination and reflectance estimation,” *IEEE Transactions on Image Processing*, vol. 24, no. 12, pp. 4965–4977, December 2015.
- [40] Y. Xu, W. Yin, Z. Wen, and Y. Zhang, “An alternating direction algorithm for matrix completion with nonnegative factors,” *Frontiers of Mathematics in China*, vol. 7, no. 2, pp. 365–384, April 2012.
- [41] J.-F. Cai, E. J. Candes, and Z. Shen, “A singular value thresholding algorithm for matrix completion,” *SIAM Journal on Optimization*, vol. 20, pp. 1956–1982, March 2010.
- [42] K. He, J. Sun, and X. Tang, “Guided image filtering,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 6, pp. 1397–1409, June 2013.
- [43] C. Lee, C. Lee, and C. Kim, “Contrast enhancement based on layered difference representation,” *IEEE International Conference on Image Processing*, pp. 965–968, September 2012.
- [44] Z. Ren, J. Li, S. Liu, and B. Zeng, “Meshflow video denoising,” *IEEE International Conference on Image Processing*, pp. 2966–2970, September 2017.
- [45] K. Gu, S. Wang, G. Zhai, S. Ma, X. Yang, W. Lin, W. Zhang, and W. Gao, “Blind quality assessment of tone-mapped images via analysis of information, naturalness, and structure,” *IEEE Transactions on Multimedia*, vol. 18, no. 3, pp. 432–443, March 2016.
- [46] K. Gu, W. Lin, G. Zhai, X. Yang, W. Zhang, and C. W. Chen, “No-reference quality metric of contrast-distorted images based on information maximization,” *IEEE Transactions on Cybernetics*, vol. 47, no. 12, pp. 4559–4565, December 2017.
- [47] K. Gu, G. Zhai, W. Lin, X. Yang, and W. Zhang, “No-reference image sharpness assessment in autoregressive parameter space,” *IEEE Transactions on Image Processing*, vol. 24, no. 10, pp. 3218–3231, October 2015.
- [48] R. Liu, L. Ma, Y. Wang, and L. Zhang, “Learning converged propagations with deep prior ensemble for image enhancement,” *IEEE Transactions on Image Processing*, vol. 28, no. 3, pp. 1528–1543, March 2019.
- [49] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, “Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising,” *IEEE Transactions on Image Processing*, vol. 26, no. 7, pp. 3142–3155, July 2017.
- [50] K. Zhang, W. Zuo, and L. Zhang, “Ffdnet: Toward a fast and flexible solution for cnn-based image denoising,” *IEEE Transactions on Image Processing*, vol. 27, no. 9, pp. 4608–4622, Sep. 2018.



Xutong Ren received her B.S. degree in computer science from Peking University, Beijing, China, in 2019. She is currently a M.S. student in machine learning at Carnegie Mellon University, Pittsburgh, PA, U.S. Her research interests include computer vision and machine learning.



Wenhan Yang (M'18) received the B.S degree and Ph.D. degree (Hons.) in computer science from Peking University, Beijing, China, in 2012 and 2018. He is currently a postdoctoral research fellow with the Department of Computer Science, City University of Hong Kong. Dr. Yang was a Visiting Scholar with the National University of Singapore, from Sep. 2015 to Sep. 2016 and from Sep. 2018 to Nov. 2018. His current research interests include deep-learning based image processing, bad weather restoration, related applications and theories.



Wen-Huang Cheng (S'05-M'08-SM'15) received the B.S. and M.S. degrees in computer science and information engineering from National Taiwan University, Taipei, Taiwan, in 2002 and 2004, respectively, where he received the Ph.D. (Hons.) degree from the Graduate Institute of Networking and Multimedia in 2008. He is a Professor with the Institute of Electronics, National Chiao Tung University (NCTU), Hsinchu, Taiwan, where he is the Founding Director with the Artificial Intelligence and Multimedia Laboratory (AIMMLab). Before joining NCTU, he led the Multimedia Computing Research Group at the Research Center for Information Technology Innovation (CITI), Academia Sinica, Taipei, Taiwan, from 2010 to 2018. His current research interests include multimedia, artificial intelligence, computer vision, machine learning, social media, and financial technology. He has received numerous research and service awards, including the 2018 MSRA Collaborative Research Award, the Outstanding Reviewer Award of 2018 IEEE ICME, the 2017 Ta-Yu Wu Memorial Award from Taiwans Ministry of Science and Technology (MOST), the 2017 Significant Research Achievements of Academia Sinica, the 2016 Y. Z. Hsu Scientific Paper Award, the Outstanding Youth Electrical Engineer Award from the Chinese Institute of Electrical Engineering in 2015, the Top 10% Paper Award from the 2015 IEEE MMSP, the K. T. Li Young Researcher Award from the ACM Taipei/Taiwan Chapter in 2014. He is APSIPA Distinguished Lecturer.



Jiaying Liu (M'10-SM'17) is currently an Associate Professor with the Wangxuan Institute of Computer Technology, Peking University. She received the Ph.D. degree (Hons.) in computer science from Peking University, Beijing China, 2010. She has authored over 100 technical articles in refereed journals and proceedings, and holds 42 granted patents. Her current research interests include multimedia signal processing, compression, and computer vision.

Dr. Liu is a Senior Member of IEEE, CSIG and CCF. She was a Visiting Scholar with the University of Southern California, Los Angeles, from 2007 to 2008. She was a Visiting Researcher with the Microsoft Research Asia in 2015 supported by the Star Track Young Faculties Award. She has served as a member of Membership Services Committee in IEEE Signal Processing Society, a member of Multimedia Systems & Applications Technical Committee (MSA TC), Visual Signal Processing and Communications Technical Committee (VSPC TC) in IEEE Circuits and Systems Society, a member of the Image, Video, and Multimedia (IVM) Technical Committee in APSIPA. She has also served as the Associate Editor of IEEE Trans. on Image Processing, and Elsevier JVCI, the Technical Program Chair of IEEE VCIP-2019/ACM ICMR-2021, the Publicity Chair of IEEE ICME-2020/ICIP-2019, and the Area Chair of CVPR-2021/ECCV-2020/ICCV-2019. She was the APSIPA Distinguished Lecturer (2016-2017).