

## Research Article

# Revisit Retinex Theory: Towards a Lightness-Aware Restorer for Underexposed Images

Lin Zhang , Anqi Zhu, Ying Shen , Shengjie Zhao , and Huijuan Zhang

School of Software Engineering, Tongji University, Shanghai, China

Correspondence should be addressed to Ying Shen; yingshen@tongji.edu.cn

Received 10 February 2020; Accepted 10 June 2020; Published 2 July 2020

Academic Editor: Daniel Zaldivar

Copyright © 2020 Lin Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We investigate how to correct exposure of underexposed images. The bottleneck of previous methods mainly lies in their naturalness and robustness when dealing with images with various exposure levels. When facing well-exposed or extremely underexposed images, they may produce over- or underenhanced outputs. In this paper, we propose a novel retinex-based approach, namely, LiAR (short for lightness-aware restorer). The word “lightness-aware” refers to that the estimated illumination not only is a component to be adjusted but also serves as a measure that reflects the brightness of the scene, determining the degree of adjustment. In this way, underexposed images can be restored adaptively according to their own brightness. Given an image, LiAR first estimates its illumination map using a specially designed loss function which can ensure the result’s color consistency and texture richness. Then adaptive correction is performed to get properly exposed output. LiAR is based on internal optimization of the single test image and does not need any prior training, implying that it can adapt itself to different settings per image. Additionally, LiAR can be easily extended to the video case due to its simplicity and stability. Experiments demonstrate that facing images/videos with various exposure levels, LiAR can achieve robust and real-time correction with high contrast and naturalness. The relevant code and collected data are publicly available at <https://cslinzhang.github.io/LiAR-Homepage/>.

## 1. Introduction

Poor lighting conditions can cause serious quality degradation of captured images and videos. For example, images taken under low-light conditions look dark overall, and back-lighting tends to cause illegible surface details in backlit region. Although the restoration of underexposed images has been a long-standing problem with a great progress made over the past decade, developing a practical effective restorer remains a challenge.

Various research studies have been done for exposure correction of underexposed images and one of the most widely used paradigms is retinex theory [1], which assumes that the sensations of color have a strong correlation with reflectance and illumination. Each color area is composed of red, green, and blue primary colors of a given wavelength, and these three primary colors determine the color of each unit area. Specifically, according to retinex theory, an image

$I$  can be decomposed into pixelwise product of reflectance  $R$  and illumination  $S$  as

$$I^c(\mathbf{x}) = R^c(\mathbf{x}) \cdot S(\mathbf{x}), \quad (1)$$

where  $c \in \{R, G, B\}$  and  $\mathbf{x}$  denotes the spatial location. It needs to be noted that for simplicity, the three color channels are usually assumed to have the same illumination [2]. According to retinex theory, an ill-exposed image is caused by its poor illumination map. Thus, the main technical challenge to restore an underexposed image is to estimate its illumination map and then to adjust it properly.

Resorting to machine learning tools to conquer the problem of illumination map estimation is a recent trend [3–5]. However, learning-based approaches have a potential drawback in their generalization capability. They rely heavily on the training data, implying that their performance may deteriorate noticeably once the conditions they were trained on are not satisfied anymore. Such a phenomenon is illustrated in Section 4.

Unlike supervised learning-based schemes, in this paper, we introduce a “zero-shot” scheme to fulfill the task of illumination map estimation. By “zero-shot,” we mean that our approach does not need any prior image examples or prior training. In our scheme, the illumination map of the given image is obtained by iteratively minimizing a specially designed loss function. Such a loss comprises two terms and they are devised to ensure color consistency and texture richness of the restored result, respectively. The illumination maps estimated in this way can endow the restored results with finer details and more natural appearances.

With the illumination map at hand, in some pipelines, it is simply removed to restore the reflectance map [6, 7]; in some other pipelines, it is adjusted further with some fixed predefined rules [8, 9]. It is easy to see that neither of the aforementioned two ways of processing illumination maps takes the brightness of the input image into full consideration. Consequently, when encountering well-exposed inputs, these methods may produce overenhanced results while for extremely underexposed inputs, their outputs are inclined to be underenhanced. In this paper, we explicitly model the impact brought by the image’s brightness and propose a simple yet effective strategy, relying on the mean brightness of the illumination map, to modify the estimated illumination map. This strategy can adaptively stretch the contrast of both bright and dark images. In this sense, we claim that our pipeline for underexposed image restoration is “lightness-aware.”

The contribution of this work is summarized as follows:

- (i) A lightness-aware restorer for underexposed images, namely, LiAR (short for lightness-aware restorer), is proposed. Its efficiency and efficacy have been quantitatively and qualitatively validated by experiments (refer to Section 4 for details).
- (ii) LiAR does not require prior training; instead, it depends on internal optimization of the single input image. Hence, LiAR has a preeminent generalization capability and can be widely applicable to various shooting scenes and kinds of illumination conditions.
- (iii) In LiAR, to optimize the illumination map of the input image, a novel loss is proposed. Such a loss can guarantee that the restored result has color consistency with the input and that it has rich texture details.
- (iv) To modify the estimated illumination map adaptively to the input image’s lightness, a strategy incorporating the mean of the illumination map is proposed and used in LiAR. This strategy allows the restored output to have appropriate brightness regardless of whether the input image is bright or dark.
- (v) LiAR can be efficiently implemented with GPU. In addition, it has excellent scalability and adaptability. Hence, it can be easily extended to enhance underexposed videos. It is worth mentioning that because of LiAR’s property of lightness awareness,

compared with the outputs of other commonly used approaches, the videos enhanced by LiAR do not have the shortcoming of flickering.

## 2. Related Work

Actually, conventional image enhancement methods such as histogram-based methods [10–14] can be explored to enhance underexposed images, but in most cases, their efficacy is quite limited. To tackle this problem more effectively, various methods specializing on this task were proposed, which fall roughly into two categories, heuristic ones and data-driven learning-based ones.

*2.1. Heuristic Methods.* Early attempts [6, 7] based on retinex theory remove the illumination and directly extract the reflectance as the enhanced results. Wang et al. [2] proposed a bright-pass filter to decompose an image into reflectance and illumination. Guo et al. [8] estimated the illumination map by imposing a structure prior on it to generate outputs with rich details. However, it neglects the color consistency, resulting in local lightness order error. In [9], Zhang et al. derived an ADMM-based procedure [15] for solving the optimization problem of illumination estimation. Despite its effectiveness in contrast enhancement, it may produce overenhancement artifacts when inputs are properly exposed images because of the fixed transformation rule used to adjust the illumination map. In addition to retinex theory, other commonly used technologies are fusion and S-curve adjustment model. Liu and Zhang [16] proposed a detail-preserving underexposed image enhancement method based on multiexposure fusion mechanism. Fusion mechanism can also be used in video enhancement, such as [17, 18]. Yuan and Sun [19] proposed an automatic exposure correction method using S-curve tone mapping. Later, the authors extended their work to correct ill-exposed videos [20]. However, the parameterized S-curve adopted in these methods may compress the midtones, and thus the output images look too flat and unnatural. Zhang et al. [21] designed a CNN (convolutional neural network) [22] to estimate the best-fitting S-curve of the input test image. To avoid loss of details in midtones, they resorted to guided filtering but this might lead to edge distortion in the output.

*2.2. Data-Driven Methods.* Recent studies on exposure correction are mostly based on machine learning. Dale et al. [23] first established a database comprising 1 million images and executed a visual search in the database. In [24], Bychkovsky et al. made a collection of 5,000 example input-output pairs that enables supervised learning. Yan et al. [25] trained deep neural networks to capture sophisticated photographic styles and modeled local adjustments that depend on image semantics. Shen et al. [26] proposed MSR-net based on multiscale retinex theory and trained it on synthesized pairwise images. In [27], Li and Wu proposed a learning-based technique of back-lit image restoration, including segmentation of back-lit and front regions and spatially adaptive tone mapping. Different from above

“black-box” models, Hu et al. [28] employed a deep reinforcement learning-based approach to provide users with an understandable solution. Based on retinex theory, Wang et al. [3] trained an illumination mapping estimation network on the new dataset they built, including underexposed images and expert-retouched references. The performance of these learning-based methods highly depends on the training dataset despite the fact that building such a dataset including various types of illumination and contents is a challenging task itself.

In this work, we take the brightness level of the input image into consideration when correcting its illumination map. Such a lightness-aware strategy can avoid overenhancement effectively. Unlike data-driven schemes, we introduce a “zero-shot” scheme to fulfill the task of illumination map estimation so that we can ensure that LiAR will perform consistently well for images spreading over a wide range of exposure levels.

### 3. Method

**3.1. General Pipeline of LiAR.** Our underexposed image restorer LiAR is established based on retinex theory (equation (1)) and accordingly, its pipeline comprises two stages, illumination estimation and exposure correction, as illustrated in Figure 1. Given an input image  $I$ , we first separate the illumination map  $S$  from  $I$  (details for illumination estimation are presented in Section 3.2) and then modify  $S$  according to its own average brightness. Finally, the restored result  $\hat{I}$  is obtained by applying the corrected illumination  $\hat{S}$  to the scene reflectance as

$$\hat{I}(\mathbf{x}) = \frac{I(\mathbf{x})}{S(\mathbf{x})} \cdot \hat{S}(\mathbf{x}), \quad (2)$$

where  $\hat{S}$  and  $\hat{I}$  are the corrected illumination map and the restored result, respectively.

In existing retinex-based methods [8, 9],  $S$  is usually adjusted using a fixed predefined rule. However, it should be noted that real inputs may have various lightness levels, such as extremely dark ones or normally exposed ones, and they actually require different levels of illumination adjustment. To this end,  $S$  has two roles in our approach, an illumination component that needs to be adjusted and also a measure that reflects the brightness of the scene, determining the degree of adjustment. Its latter role accounts for the “lightness awareness” of LiAR. Inspired by gamma transformation in image-tone mapping, our lightness-aware illumination adjustment scheme is designed as

$$\hat{S}(\mathbf{x}) = [S(\mathbf{x})]^{\bar{S}}, \quad (3)$$

where  $\bar{S}$  is the mean brightness of the illumination map  $S$ , serving as a measure that reflects the brightness of the scene. Using this transformation, the degree of adjustment can be determined by the illumination brightness. For example, originally darker images with  $\bar{S}$  close to 0 will be greatly enhanced, while well-exposed images with higher  $\bar{S}$  will remain as they are.

Several examples are shown in Figure 2 to demonstrate the capability of LiAR. In the first row of Figure 2,  $I_1 \sim I_3$  are

three input images and  $S_1 \sim S_3$  are their estimated illumination maps. Using LiAR, the corresponding restoration results  $\hat{I}_1 \sim \hat{I}_3$  along with their corrected illumination maps  $\hat{S}_1 \sim \hat{S}_3$  are obtained and shown in the second row of Figure 2. It can be observed that with our lightness-aware strategy, the illumination maps can be adaptively adjusted.

Next, we will discuss how to estimate the illumination map from a given input image.

**3.2. Illumination Estimation.** Given an image  $I$ , its illumination map is expected to be estimated in such a way that the final restored output should have the color consistency with the input and have rich textures. In LiAR, these two goals are achieved by imposing two constraints on illumination map optimization, one for color consistency and one for texture richness.

**3.2.1. Color Consistency Loss.** When an image is processed, its intensities of pixels are normalized to  $[0, 1]$ . For each color channel, according to equations (2) and (3), the restored intensity at position  $\mathbf{x}$  can be written as

$$\hat{I}^c(\mathbf{x}) = \frac{I^c(\mathbf{x})}{S(\mathbf{x})} \cdot [S(\mathbf{x})]^{\bar{S}}, \quad c \in \{R, G, B\}. \quad (4)$$

When the restored intensity in one channel  $\hat{I}^c(\mathbf{x})$  overflows, which means  $\hat{I}^c(\mathbf{x}) > 1$ , to ensure that the restored intensities fall in  $[0, 1]$ ,  $\hat{I}^c(\mathbf{x})$  will be cut off to  $\hat{I}^c(\mathbf{x}) = 1$ . In this situation, the color consistency between the input image and the output will be broken since

$$\left( \hat{I}^R(\mathbf{x}), \hat{I}^G(\mathbf{x}), \hat{I}^B(\mathbf{x}) \right) \not\parallel \left( I^R(\mathbf{x}), I^G(\mathbf{x}), I^B(\mathbf{x}) \right), \quad (5)$$

where  $\mathbf{a} \not\parallel \mathbf{b}$  means that the vectors  $\mathbf{a}$  and  $\mathbf{b}$  are not parallel to each other. To avoid this,  $S(\mathbf{x})$  should be

$$S(\mathbf{x}) \geq \max_{c \in \{R, G, B\}} I^c(\mathbf{x}). \quad (6)$$

In order to consider the color constraint and other constraints together in optimization, equation (6) is expressed as a loss term  $L_c$  (short for  $L_{\text{color}}$ ):

$$L_c = \sum_{\mathbf{x} \in \Omega} \text{relu} \left( \max_{c \in \{R, G, B\}} I^c(\mathbf{x}) - S(\mathbf{x}) \right), \quad (7)$$

$$\text{relu}(\mathbf{x}) = \begin{cases} 0, & \mathbf{x} < 0, \\ \mathbf{x}, & \mathbf{x} \geq 0. \end{cases} \quad (8)$$

From the definition of  $L_c$  in equation (7), it is easy to know that only when  $S(\mathbf{x})$  is smaller than  $\max_{c \in \{R, G, B\}} I^c(\mathbf{x})$ ,  $L_c$  will contribute to the loss. In our implementation,  $\max_{c \in \{R, G, B\}} I^c(\mathbf{x})$  is chosen as the initial estimation of the illumination map  $S_0(\mathbf{x})$ .

**3.2.2. Texture Richness Loss.** In an image, usually the illumination intensity of a surface is relatively flat, and the contrast of the surface should be enhanced to ensure texture richness. If the estimated illumination of a surface spatially fluctuates as texture changes, the calculated reflectance of the

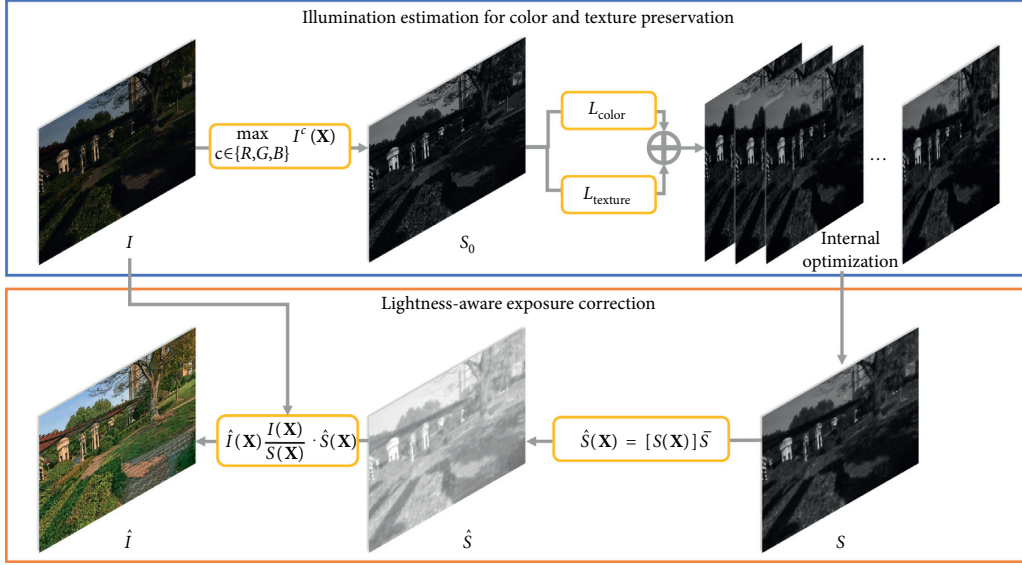


FIGURE 1: Pipeline of our lightness-aware underexposed image restorer, LiAR. For the input image  $I$ , the initial estimation of its illumination map  $S_0$  is taken as the maximum of  $\{R, G, B\}$  channels' intensities. Then, the illumination map  $S$  of the scene is estimated from  $S_0$  by internal optimization under color and texture constraints  $L_{\text{color}}$  and  $L_{\text{texture}}$ . After that, lightness-aware illumination adjustment is applied to  $S$  to get the corrected illumination map  $\hat{S}$ . The final restoration result  $\hat{I}$  is obtained as the production of the reflectance and the corrected illumination  $\hat{S}$ .

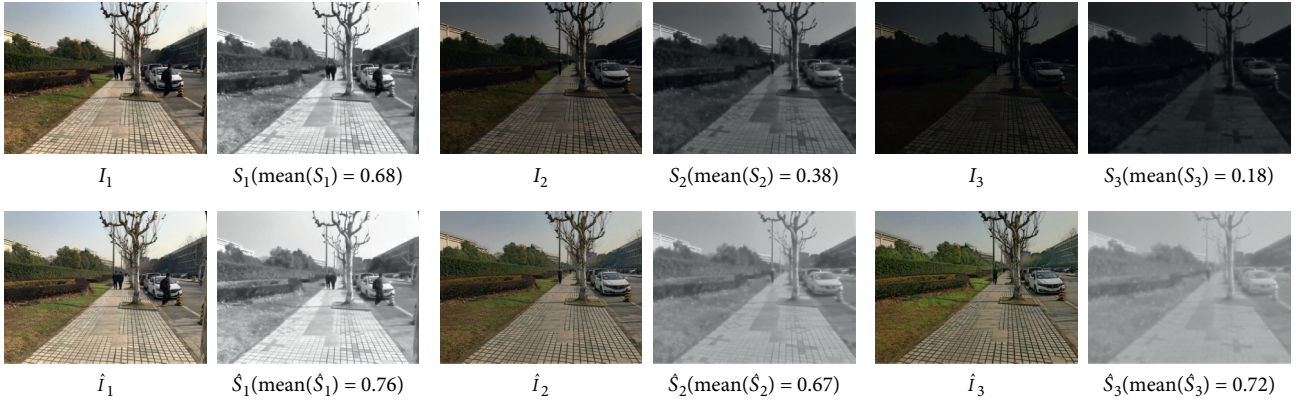


FIGURE 2:  $I_1 \sim I_3$  are three input images with different exposure levels.  $S_1 \sim S_3$  are the estimated illumination maps of  $I_1 \sim I_3$ , respectively, and  $\hat{S}_1 \sim \hat{S}_3$  are the corresponding corrected illumination maps.  $\hat{I}_1 \sim \hat{I}_3$  are the outputs of LiAR by taking  $I_1 \sim I_3$  as inputs.

scene will be flatter than the ground truth, resulting in smoothed texture in the output. Therefore, in order to ensure that the texture is enhanced, it is necessary to make the illumination intensity as smooth as possible, which can be expressed as a loss term  $L_t$  (short for  $L_{\text{texture}}$ ) of illumination estimation:

$$L_t = \sum_{\mathbf{x} \in \Omega} w_x(\mathbf{x})(\partial_x S(\mathbf{x}))^2 + w_y(\mathbf{x})(\partial_y S(\mathbf{x}))^2 + \lambda_t (S(\mathbf{x}) - S_0(\mathbf{x}))^2, \quad (9)$$

where  $w(\mathbf{x})$  is the weight at each pixel and  $x$  and  $y$  represent the horizontal and vertical directions, respectively.  $\lambda_t$  is a predefined parameter. The term  $\lambda_t (S(\mathbf{x}) - S_0(\mathbf{x}))^2$  is used to control the estimation result not to deviate too much from the initial estimation. The remaining key issue is how to

design the weight  $w_x(\mathbf{x})$  and  $w_y(\mathbf{x})$ . Note that a region with small gradients usually corresponds to a flat surface in the scene and needs to be smoothed. Inspired by RTV loss [29], a simplified weight, inversely proportional to the gradient, is designed as

$$w_x(\mathbf{x}) = \frac{1}{G \circ (\partial_x I_g(\mathbf{x}))^2}, \quad (10)$$

where  $G$  is a Gaussian filter and  $I_g$  is the greyscale map of the input image.  $w_y$  can be computed in a similar way. The weight terms only need to be computed once at the beginning of processing.

Combining the two loss terms  $L_c$  and  $L_t$  via a parameter  $\lambda$ , we get the loss function of the illumination estimation:



$$L = L_t + \lambda L_c. \quad (11)$$

At this point, given an image, its illumination map can be estimated by iteratively minimizing  $L$ .

**3.3. Implementation Details.** LiAR is implemented with PyTorch [30]. Images are converted into tensors for parallel computing. All the operations involved in computing  $L$  are differentiable. With respect to the optimization algorithm for updating the illumination map, the SGD (stochastic gradient descent) algorithm [31] is used.

There are two hyperparameters in LiAR,  $\lambda_t$  and  $\lambda$ . In order to keep the values of the three terms of equation (9) in the same order of magnitude, we set  $\lambda_t = 0.1$ .  $\lambda$  is designed to control the weight of two losses, especially the color consistency loss  $L_c$ . The value of  $L_c$  can reflect the color distortion of the corrected image. Thus, we need to set the value of  $\lambda$  high enough to make sure that there is no noticeable color distortion. In all experiments, we set  $\lambda = 100$ , which is high enough to make the color consistency term as a strong constraint.

## 4. Experiments

We conducted experiments on real-world images to compare the performance of LiAR with the state-of-the-art or representative approaches for underexposed image restoration. Furthermore, the ablation study is performed to evaluate the impact of each component of LiAR. Additionally, we applied LiAR to enhance underexposed videos and then compared its results with other competitors in this field.

All the experiments were carried out on a workstation with a 3.0 GHz Intel Core i7-5960X CPU and an Nvidia GeForce GTX 980Ti GPU.

### 4.1. Evaluation on Underexposed Images

**4.1.1. Datasets.** Since our goal is to evaluate the capability of restoration on different exposure levels, the dataset is desired to contain images with various exposure levels. To this end, the experiments were performed on 1,500 real-world images taken from  $IE_{ps}D$  [32], which was established for studying the problem of exposure level assessment. We partitioned these images into three groups, 500 images for each group, according to their exposure settings. Three groups are “well exposed” (Group A), “slightly underexposed” (Group B), and “severely underexposed” (Group C).

**4.1.2. Compared Methods.** LiAR was compared with eight underexposed image restorers, including (1) HE [12], (2) CLAHE [14], (3) Retinex [6], (4) Yuan and Sun’s method [19], (5) LIME [8], (6) Exposure [28], (7) DeepUPE [3], and (8) ExCNet [21].

**4.1.3. Objective Evaluation.** The performance of underexposed image restoration methods was evaluated with two objective metrics, CDIQA (contrast-distorted image quality assessment) [33] and LOE (lightness order error) [2].

CDIQA is a no-reference quality assessment of contrast-distorted images, which can be considered as a metric for richness of image details. A higher CDIQA value roughly corresponds to higher contrast. LOE is a measure to objectively assess the naturalness preservation between the input and enhanced output. Ideally, if the enhancement approach does not violate the relative lightness order of pixel values in the input image, the associated LOE measure would be zero. Thus, a lower LOE value roughly corresponds to less artifacts caused by restoration.

The results over 1,500 test images are reported in Table 1. It can be seen that for every case, LiAR can obtain a high CDIQA value and a low LOE value, demonstrating its superiority in restoring the input image’s details while keeping its naturalness. It also corroborates that LiAR has a strong generalization capability and can be employed to cope with images spreading over a wide range of exposure levels. By contrast, the performance and robustness of the competitors are apparently inferior to LiAR. For example, though Exposure [28] and DeepUPE [3] perform quite well when dealing with well-exposed images (Group A), their performance deteriorates significantly on obviously underexposed ones (Groups B and C). As for LIME [8] and ExCNet [21], for all cases, they can achieve high CDIQA values, indicating that their outputs are of high contrast. However, their LOE values are also quite large, implying that they suffer from the problem of overenhancement.

**4.1.4. Visual Quality.** Figure 3 compares the restoration results of the competing methods on a severely underexposed input. It can be seen that facing such an extremely dark image, the results of the learning-based methods Exposure [28] and DeepUPE [3] look quite dim and the details are invisible. Figure 4 shows the restoration results on a slightly ill-exposed image. For this case, S-curve-based approaches [19, 21] tend to get flat results, meaning that midtone textures are significantly compressed. In both Figures 3 and 4, the results of LIME [8] obviously suffer from the unwanted artifacts. By contrast, the outputs of LiAR are natural and of high contrast. These observations are consistent with the quantitative evaluations reported in Table 1.

**4.1.5. Results on Different Exposure Levels.** In order to demonstrate the generalization capability of LiAR and the drawback of learning-based approaches, we conducted experiments on images with different exposure levels. We compare LiAR with a state-of-the-art learning-based method DeepUPE [3]. As shown in Figure 5, DeepUPE [3] fails to enhance severely underexposed images. The underlying reason is that its training dataset does not cover the extremely underexposed cases like Figures 5(b) and 5(c), which shows that learning-based approaches rely heavily on the training data, implying that their performance may deteriorate noticeably once the conditions they were trained on are not satisfied anymore. By contrast, our proposed approach LiAR, as an image-specific method, performs consistently well for images spreading over a wide range of exposure levels.

TABLE 1: Objective evaluation results of the competing methods.

Methods	Pretrained	Group A		Group B		Group C		All Groups	
		CDIQA	LOE	CDIQA	LOE	CDIQA	LOE	CDIQA	LOE
HE [12]	NO	2.7240	0.0540	2.9697	0.0478	3.0058	0.0537	2.8998	0.0518
CLAHE [14]	NO	3.2602	0.1352	2.8724	0.0988	2.1903	0.0632	2.7743	0.0991
Retinex [6]	NO	3.1847	0.1566	3.1693	0.1625	3.0282	0.1943	3.1274	0.1711
Yuan and Sun [19]	NO	2.8227	0.3041	3.1804	0.1105	2.3889	0.0354	2.7973	0.1500
LIME [8]	NO	3.2417	0.1955	3.3036	0.1490	2.9276	0.1352	3.1576	0.1599
Exposure [28]	YES	3.2930	0.0295	2.7563	0.0298	2.3231	0.0606	2.7908	0.0400
DeepUPE [3]	YES	3.3312	0.0680	2.7525	0.0294	2.1580	0.0254	2.7472	0.0409
ExCNet [21]	NO	3.0677	0.2893	3.2757	0.1239	2.6697	0.0566	3.0044	0.1566
<b>LiAR</b>	<b>NO</b>	<b>3.2798</b>	<b>0.0324</b>	<b>3.2750</b>	<b>0.0576</b>	<b>2.7801</b>	<b>0.0569</b>	<b>3.1116</b>	<b>0.0490</b>

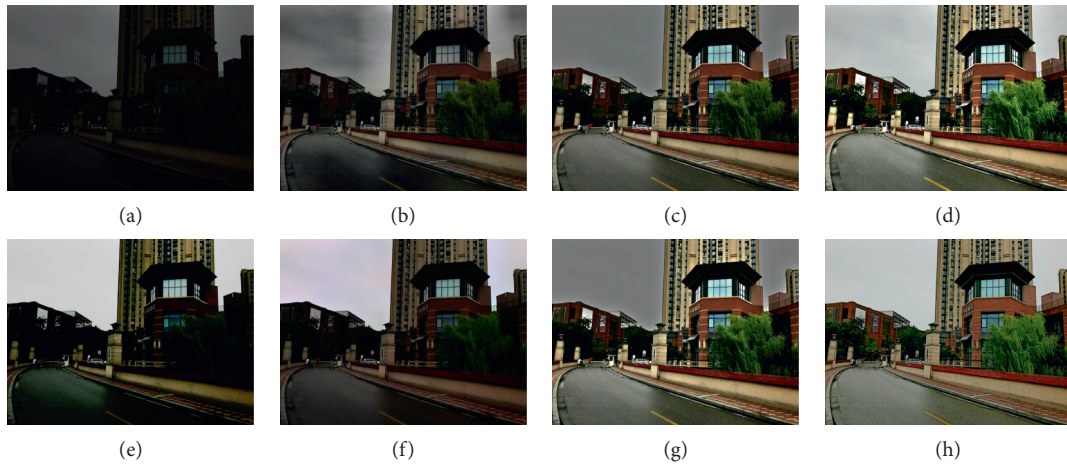


FIGURE 3: Comparison on a severely underexposed image. (a) Input. (b) CLAHE [6]. (c) Yuan and Sun [19]. (d) LIME [8]. (e) Exposure [28]. (f) DeepUPE [3]. (g) ExCNet [21]. (h) LiAR.

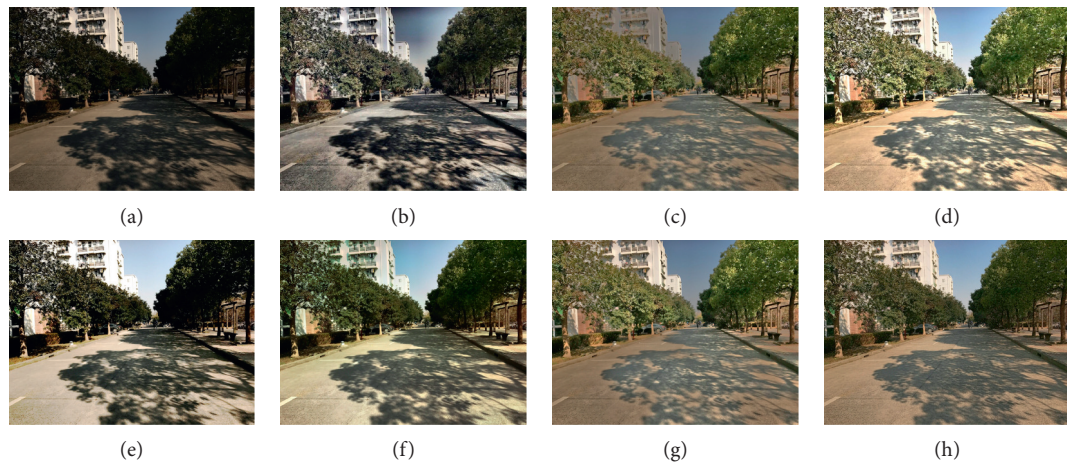


FIGURE 4: Comparison on a slightly underexposed image. (a) Input. (b) CLAHE [6]. (c) Yuan and Sun [19]. (d) LIME [8]. (e) Exposure [28]. (f) DeepUPE [3]. (g) ExCNet [21]. (h) LiAR.

**4.1.6. Ablation Study.** We performed an ablation study to analyze the importance of each component of LiAR, and the results are summarized in Table 2.

The first three settings in the table correspond to removing two loss terms and the lightness-aware design on the basis of LiAR. It can be seen that removing  $L_c$  can achieve

high contrast while causing serious artifacts. On the contrary, removing  $L_t$  and only using  $L_c$  lead to low contrast while keeping the lightness order consistency. Therefore, it can be confirmed that combining  $L_c$  and  $L_t$  can help to balance contrast and fidelity of the restored results. If the lightness-aware illumination correction strategy is replaced



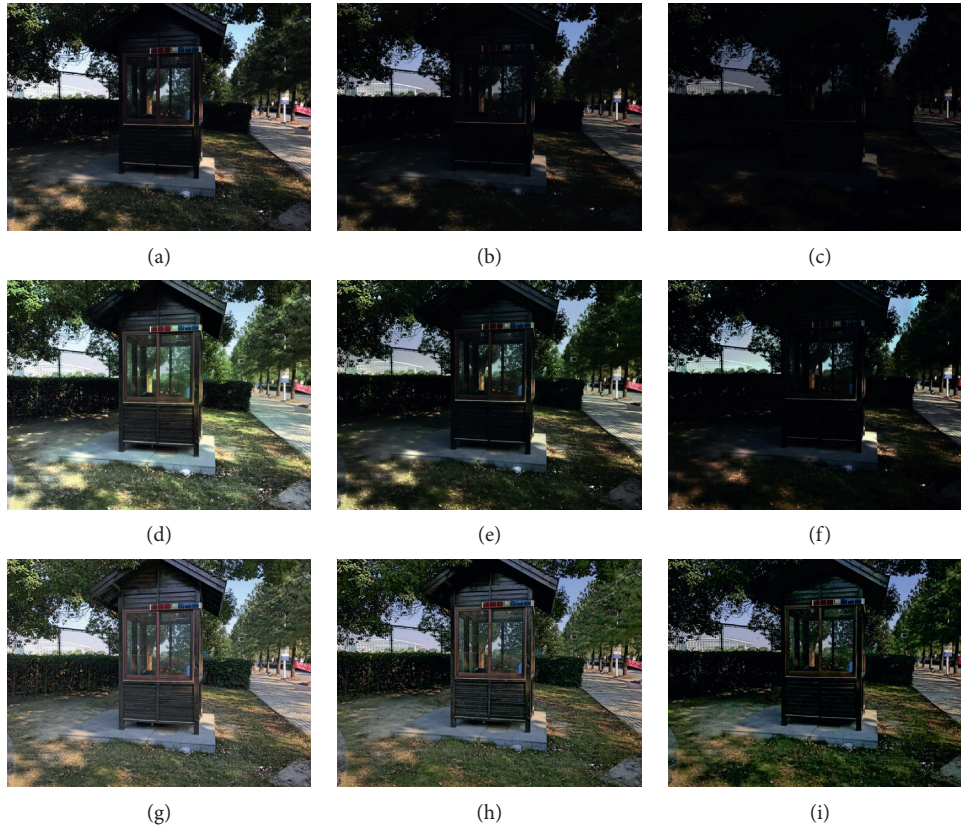


FIGURE 5: (a–c) Three underexposed photos shot in the same scene with different exposure levels. (a) Slightly underexposed. (b, c) Relatively severely underexposed. (d–f) Restoration results of a state-of-the-art method DeepUPE [3], which is a learning-based approach. (g–i) Results of LiAR, the approach proposed in this paper. It is not difficult to find that DeepUPE [3] can restore slightly underexposed images very well while performing quite poor for the severely underexposed ones. By contrast, LiAR, as an image-specific approach, can deal with images with a wide range of exposure levels quite well.

TABLE 2: Ablation study.

Settings	Metrics	Group A	Group B	Group C	All
No $L_c$	CDIQA	3.3075	3.2836	3.0467	3.2126
	LOE	0.2475	0.2955	0.2407	0.2613
No $L_t$	CDIQA	3.1739	3.1447	2.6133	2.9773
	LOE	0.0196	0.0214	0.0264	0.0225
No lightness-aware $\gamma = 0.4$	CDIQA	3.2868	2.8166	2.3288	2.8107
	LOE	0.0269	0.0360	0.0435	0.0355
LiAR ( $\gamma = \bar{S}$ )	CDIQA	<b>3.2798</b>	<b>3.2750</b>	<b>2.7801</b>	<b>3.1116</b>
	LOE	<b>0.0324</b>	<b>0.0576</b>	<b>0.0569</b>	<b>0.0490</b>

with a fixed gamma transformation like [8, 9] where  $\gamma = 0.4$ , the performance on Group A is satisfied while the performance on Groups B and C is much inferior to LiAR. The underlying reason is that the fixed rule cannot adaptively adjust the illumination maps.

**4.2. Evaluation on Underexposed Videos.** Though LiAR is initially designed for coping with a single image, it can be easily adapted to the video case. In this experiment, its performance for underexposed video restoration was evaluated.

**4.2.1. Dataset and Compared Methods.** Since there is no publicly available dataset for the study of underexposed video restoration, we collected such a dataset by ourselves which includes 112 video clips with back-lighting or low-light illumination conditions. They were also classified into three groups, “well exposed” (Group A, 32 clips), “slightly underexposed” (Group B, 42 clips), and “severely underexposed” (Group C, 38 clips). We compared LiAR with four representative approaches in this field, including (1) virtual exposure [34], (2) Dong et al.’s method [35], (3) the traditional image enhancement method HE [12], and (4) ExCNet [21].

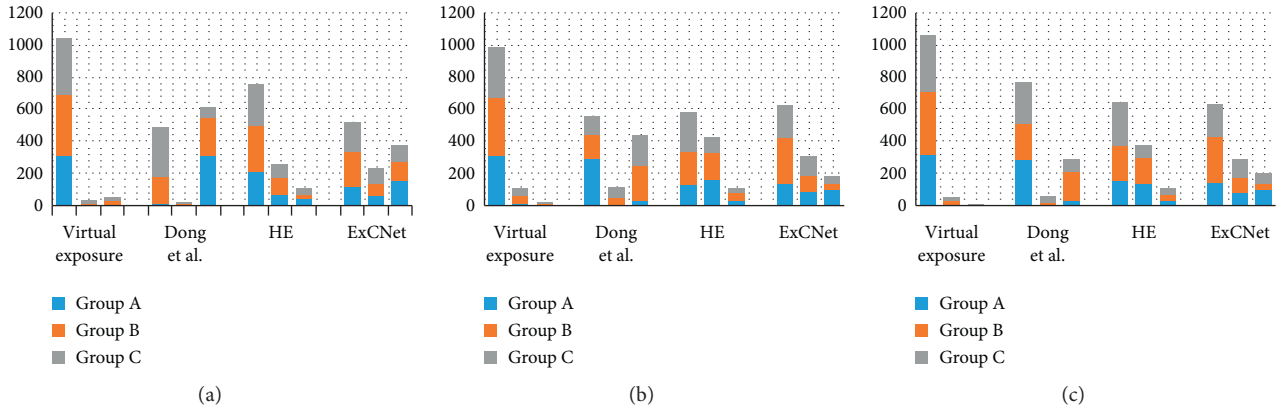


FIGURE 6: Results of pairwise comparison user study. The pairwise comparison was conducted between the result of LiAR and the result of one of the four competing methods. There are three bars in each pairwise comparison, which are the number of the votes for “our method,” “competitor,” and “no preference” from left to right. (a) Details visibility. (b) Visual naturalness. (c) Overall preference.

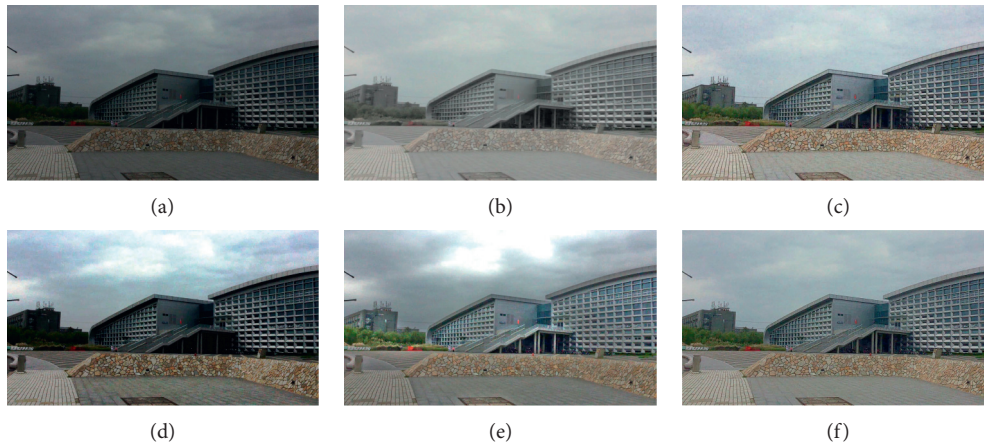


FIGURE 7: (a) A frame taken from a video clip. Restoration results of (a) by using (b) virtual exposure [34], (c) Dong et al.’s method [35], (d) HE [12], (e) ExCNet [21], and (f) LiAR, respectively.

**4.2.2. Pairwise Comparison User Study.** We conducted a user study with ten volunteers (5 males and 5 females) to make pairwise comparison between the corrected results of our method and those of the compared methods. This comparison was made from three aspects, including “details visibility,” “visual naturalness,” and “overall preference.” For each pairwise comparison, the group of videos and the order of method pairs were randomized to avoid subjective bias. There were three options for the user to choose: “left is better,” “right is better,” or “no preference.”

The results of the user study are summarized in Figure 6. There are three bars in each pairwise comparison corresponding to the subject’s preference, which are the number of the votes for “our method,” “competitor,” and “no preference” from left to right. The number of videos from different groups is represented with different colors. The results in Figure 6 clearly demonstrate that no matter which criterion is used, the participants showed a strong bias in preference towards the correction results of LiAR.

TABLE 3: The ave\_SRCC values of the evaluated approaches.

Methods	ave_SRCC
Virtual exposure [34]	0.7157
Dong et al. [35]	0.7308
HE [12]	0.0757
ExCNet [21]	0.3699
LiAR	0.7198

**4.2.3. Visual Quality.** Figure 7(a) is the input frame while Figures 7(b) ~ 7(f) are the restoration results of the competing methods. It can be observed that the result of LiAR has better color consistency, finer details, and less over-enhancement artifacts.

Unlike the case of processing a single image, when restoring an underexposed video, in addition to ensuring the restoration quality of each frame, we must ensure the smoothness of the video content, that is, we cannot introduce flickering artifacts during the restoration process. Therefore, the restoration algorithm is expected to have the ability to



TABLE 4: The average time cost (ms) for processing one frame.

Methods	Device	1080P	720P	480P
Virtual exposure [34]	CPU	4175	1877	702
Dong et al. [35]	CPU	2008	675	228
HE [12]	CPU	53	28	15
ExCNet [21]	GPU	4185	3897	3823
LiAR	GPU	64	37	31

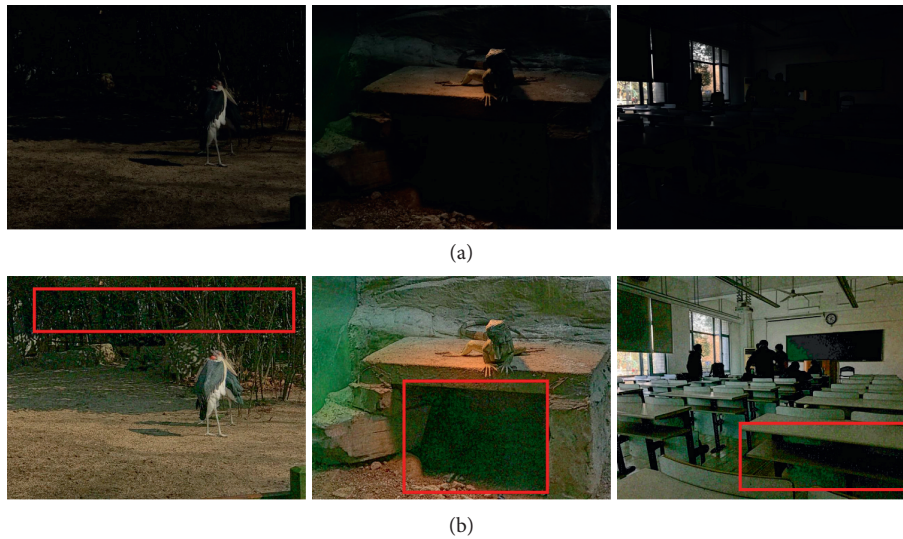


FIGURE 8: Failure cases. Extremely dark input images with noise (a) and their correction results (b).

maintain the brightness order of video frames. In this paper, to quantify the algorithm's ability to keep the brightness order of video frames, a metric "ave\_SRCC" is designed as follows.

Suppose that  $v_i$  is a video clip having  $n$  frames and the set of its average frame brightness is denoted by  $\{b_i^j\}_{j=1}^n$ . Denote by  $\hat{v}_i$  the restoration result of  $v_i$  and its set of average frame brightness is denoted by  $\{\hat{b}_i^j\}_{j=1}^n$ . Then, the ave\_SRCC is defined as

$$\text{ave\_SRCC} = \frac{1}{m} \sum_{i=1}^m \text{SRCC} \left( \{b_i^j\}_{j=1}^n, \{\hat{b}_i^j\}_{j=1}^n \right), \quad (12)$$

where  $m$  is the number of video clips and SRCC computes the Spearman rank-order correlation coefficient of two vectors [36].

ave\_SRCC values of the competing methods are listed in Table 3. It can be seen that ave\_SRCC values of virtual exposure [34], Dong et al.'s method [35], and LiAR are much higher than those of HE and ExCNet [21]. It indicates that the restored videos of the former three approaches have much less flickering artifacts than those of the latter two approaches. This conclusion is consistent with the intuitive observation when comparing results subjectively.

**4.2.4. Time Cost.** In this experiment, the running speeds of evaluated approaches are analyzed. In Table 4, the time cost of each competing method for processing one frame is presented for reference. We tested three commonly

encountered video resolutions, 1080P, 720P, and 480P. Whether the implementation was based on CPU or GPU is also reported in Table 4. It needs to be noted that for competing methods by other authors, we used their own or official implementations, and thus for virtual exposure [34] and Dong et al.'s method [35], we did not have their GPU-based implementations. LiAR's implementation is based on GPU and it consumes about 30 ms to process one  $640 \times 480$  video frame.

Figure 8 presents three examples where LiAR fails to produce visually compelling results. For the extremely dark input images with noise, we amplified the noise in the dark regions when we greatly brightened images. This is because the images collected in a dim environment usually contain noise in dark regions, and the noise will be regarded as texture information and then amplified when illumination is brightened.

## 5. Conclusions

This paper proposes LiAR, a two-phase approach for underexposed image restoration. Given an input image, LiAR first estimates its illumination map by minimizing a loss which comprises two terms used to ensure color consistency and texture richness of the output, respectively. Then, it adjusts the illumination map in a lightness-aware way. Experimental results demonstrate that images enhanced by LiAR own high contrast while keeping naturalness. In addition, LiAR can be easily extended to the video

case. Compared with other competitors for underexposed video restoration, LiAR can output frames with pleasing quality. More importantly, it can keep the brightness order among video frames quite well, which makes it avoid the flickering artifacts usually existing in outputs of other evaluated approaches.

Our future work is to design a denoising module to suppress noise in extremely dark regions. A direction is to perform denoising on the reflectance component obtained by retinex decomposition.

## Data Availability

The source code and video dataset have been made publicly available at <https://cslinzhang.github.io/LiAR-Homepage/>.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This research was funded in part by the National Natural Science Foundation of China under grant nos. 61973235, 61672380, 61972285, and 61936014 and in part by the Natural Science Foundation of Shanghai under grant no. 19ZR1461300.

## References

- [1] E. H. Land and J. J. McCann, "Lightness and retinex theory," *Journal of the Optical Society of America*, vol. 61, no. 1, pp. 1–11, 1971.
- [2] 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, 2013.
- [3] R. Wang, Q. Zhang, C.-W. Fu, X. Shen, W.-S. Zheng, and J. Jia, "Underexposed photo enhancement using deep illumination estimation," in *Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA, June 2019.
- [4] W. Chen, W. Wang, W. Yang, and J. Liu, "Deep retinex decomposition for low-light enhancement," in *Proceedings of the BMVC*, Newcastle Upon Tyne, UK, September 2018.
- [5] Y. Zhang, J. Zhang, and X. Guo, "Kindling the darkness: a practical low-light image enhancer," in *Proceedings of the 27th ACM International Conference on Multimedia*, Nice, France, October 2019.
- [6] D. J. Jobson, Z. Rahman, and G. A. Woodell, "A multiscale retinex for bridging the gap between color images and the human observation of scenes," *IEEE Transactions on Image Processing*, vol. 6, no. 7, pp. 965–976, 1997.
- [7] Z. Rahman, D. J. Jobson, and G. A. Woodell, "Multi-scale retinex for color image enhancement," in *Proceedings of the 3rd IEEE International Conference on Image Processing*, Lausanne, Switzerland, September 1996.
- [8] 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, 2017.
- [9] Q. Zhang, G. Yuan, C. Xiao, L. Zhu, and W.-S. Zheng, "High-quality exposure correction of underexposed photos," in *Proceedings of the 2018 ACM Multimedia Conference on Multimedia Conference—MM'18*, Yokohama, Japan, June 2018.
- [10] M. Abdullah-Al-Wadud, M. Kabir, M. Akber Dewan, and O. Chae, "A dynamic histogram equalization for image contrast enhancement," *IEEE Transactions on Consumer Electronics*, vol. 53, no. 2, pp. 593–600, 2007.
- [11] S. M. Pizer, E. P. Amburn, J. D. Austin et al., "Adaptive histogram equalization and its variations," *Computer Vision, Graphics, and Image Processing*, vol. 39, no. 3, pp. 355–368, 1987.
- [12] J. C. Russ, *The Image Processing Handbook*, CRC Press, Boca Raton, FL, USA, 2016.
- [13] C. Wang and Z. Ye, "Brightness preserving histogram equalization with maximum entropy: a variational perspective," *IEEE Transactions on Consumer Electronics*, vol. 51, no. 4, pp. 1326–1334, 2005.
- [14] K. Zuiderveld, "Contrast limited adaptive histogram equalization," in *Graphics Gems IV*, Elsevier, Amsterdam, Netherlands, 1994.
- [15] S. Boyd, N. Parikh, E. Chu, P. Borja, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.
- [16] S. Liu and Y. Zhang, "Detail-preserving underexposed image enhancement via optimal weighted multi-exposure fusion," *IEEE Transactions on Consumer Electronics*, vol. 65, no. 3, pp. 303–311, 2019.
- [17] Q. Zhang, Y. Nie, L. Zhang, and C. Xiao, "Underexposed video enhancement via perception-driven progressive fusion," *IEEE Transactions on Visualization and Computer Graphics*, vol. 22, no. 6, pp. 1773–1785, 2015.
- [18] Y. Zhang and S. Liu, "Non-uniform illumination video enhancement based on zone system and fusion," in *Proceedings of the 2018 24th International Conference on Pattern Recognition (ICPR)*, Beijing, China, August 2018.
- [19] L. Yuan and J. Sun, "Automatic exposure correction of consumer photographs," in *Computer Vision—CCV 2012*, Springer, Berlin, Germany, 2012.
- [20] X. Dong, Y. Lu, W. Li, and A. L. Yuille, "Temporally consistent region-based video exposure correction," in *Proceedings of the 2015 IEEE International Conference on Multimedia and Expo (ICME)*, Turin, Italy, June 2015.
- [21] L. Zhang, L. Zhang, X. Liu, Y. Shen, S. Zhang, and S. Zhao, "Zero-shot restoration of back-lit images using deep internal learning," in *Proceedings of the 27th ACM International Conference on Multimedia*, Nice, France, October 2019.
- [22] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [23] K. Dale, M. K. Johnson, K. Sunkavalli, W. Matusik, and H. Pfister, "Image restoration using online photo collections," in *Proceedings of the 2009 IEEE 12th International Conference on Computer Vision*, Kyoto, Japan, September 2009.
- [24] V. Bychkovsky, S. Paris, E. Chan, and F. Durand, "Learning photographic global tonal adjustment with a database of input/output image pairs," in *Proceedings of the CVPR*, Providence, RI, USA, June 2011.
- [25] Z. Yan, H. Zhang, B. Wang, S. Paris, and Y. Yu, "Automatic photo adjustment using deep neural networks," *ACM Transactions on Graphics*, vol. 35, no. 2, pp. 1–15, 2016.
- [26] L. Shen, Z. Yue, F. Fan, Q. Chen, S. Liu, and J. Ma, "MSR-net: low-light image enhancement using deep convolutional network," 2017, <https://arxiv.org/abs/1711.02488>.
- [27] Z. Li and X. Wu, "Learning-based restoration of backlit images," *IEEE Transactions on Image Processing*, vol. 27, no. 2, pp. 976–986, 2017.

- [28] Y. Hu, H. He, C. Xu, B. Wang, and S. Lin, "Exposure: a white-box photo post-processing framework," *ACM Transactions on Graphics*, vol. 37, no. 2, pp. 26–31, 2018.
- [29] L. Xu, Q. Yan, Y. Xia, and J. Jia, "Structure extraction from texture via relative total variation," *ACM Transactions on Graphics*, vol. 31, no. 6, pp. 139–141, 2012.
- [30] P. Adam, S. Gross, S. Chintala et al., "Automatic differentiation in pytorch," 2017.
- [31] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, UK, 2004.
- [32] L. Zhang, L. Zhang, X. Liu, Y. Shen, and D. Wang, "Image exposure assessment: a benchmark and a deep convolutional neural networks based model," in *Proceedings of the 2018 IEEE International Conference on Multimedia and Expo (ICME)*, San Diego, CA, USA, July 2018.
- [33] Y. Fang, K. Ma, Z. Wang, W. Lin, Z. Fang, and G. Zhai, "No-reference quality assessment of contrast-distorted images based on natural scene statistics," *IEEE Signal Processing Letters*, vol. 22, no. 7, pp. 838–842, 2014.
- [34] E. P. Bennett and L. McMillan, "Video enhancement using per-pixel virtual exposures," *ACM Transactions on Graphics*, vol. 24, no. 3, pp. 845–852, 2005.
- [35] X. Dong, G. Wang, Yi Pang et al., "Fast efficient algorithm for enhancement of low lighting video," in *Proceedings of the 2011 IEEE International Conference on Multimedia and Expo*, Barcelona, Spain, July 2011.
- [36] C. Spearman, "'Footrule' for measuring correlation," *British Journal of Psychology, 1904–1920*, vol. 2, no. 1, pp. 89–108, 1906.