



# Perceptive low-light image enhancement via multi-layer illumination decomposition model

Yahong Wu<sup>1</sup> · Jieying Zheng<sup>1</sup> · Wanru Song<sup>1</sup> · Feng Liu<sup>1</sup>

Received: 12 January 2021 / Revised: 1 April 2021 / Accepted: 10 April 2022 /

Published online: 14 May 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

## Abstract

The images captured under low light conditions generally have less than satisfactory visual quality. To address this issue, many low-light image enhancement methods have been studied. However, these existing algorithms mostly suffer from unnaturalness, over-enhancement and artifacts. In this paper, a perceptive low-light image enhancement via multi-layer illumination decomposition model is proposed, to preserve the naturalness and improve the contrast for low-light images. First, the contrast of the target image is defined from global, local and the effect of noise aspects. Then, inspired by the human visual system, the perceptive contrast is designed by combining the defined contrast with just-noticeable-difference transformation. Last and most importantly, the target image is decomposed in a multi-layer way based on the multi-scale adaptive filter, which utilizes the perceptive contrast to decide the variance adaptively. This step can effectively obtain multiple illumination and reflectance layers. Combining these reflectance with adjusted illumination components can generate the final enhanced result. The proposed method has better no-reference quantitative measurement results than other compared methods. Experimental results on several public challenging low-light image datasets demonstrate that the proposed method can achieve great performance in balancing the contrast, brightness and naturalness.

**Keywords** Low-light image enhancement · Perceptive contrast · Multi-layer illumination decomposition · Naturalness preservation · Human visual perception

---

This work was supported in part by National Natural Science Foundation of China under Grant 61702278, in part by Priority Academic Program Development of Jiangsu Higher Education Institutions and in part by Postgraduate Research & Practice Innovation Program of Jiangsu Province KYCX18\_0902.

---

✉ Feng Liu  
liuf@njupt.edu.cn

Yahong Wu  
wuyahonghaha@163.com

<sup>1</sup> Nanjing University of Posts and Telecommunications, No.66 Xin Mofan RD, Nanjing, 210003, China

# 1 Introduction

With the development of image technology, massive images are constantly produced by mobile phones and cameras. However, some images are of poor quality, such as the images captured under low light conditions, which may cause inconvenience to people's lives. Low visibility hinders many real applications, such as objection detection [48] and tracking [38], image matting [35], and person re-identification [28], which are often used in traffic and criminal investigations. Therefore, enhancing low-light images has a significant effect on our daily work and life. Low-light images have low dynamic range and suffer from heavy noise. For low-light image enhancement, it is a challenging task to balance the relationship among brightness, contrast and naturalness. Inspired by the above problem, many advanced low-light image enhancement methods have been proposed in the past decades.

Histogram equalization (HE) [1] is the simplest and most intuitive method, which can stretch the dynamic range of the observed image. Many extended HE methods have been developed [34, 39]. Nevertheless, these methods exist color distortion and often neglect the effect of noise, especially the noise hidden in dark areas. Inspired by an observation that an inverted low-light image looks like a haze image [53], dehazing-based methods have been proposed to enhance low-light images [10, 26]. Although dehazing-based methods can generate relatively reasonable results, a convincing physical explanation of their basic model has not been provided. With the development of machine learning and deep learning, many low-light image enhancement networks have been proposed [42, 47, 50, 52, 55]. These methods require a mass of sample pairs: low-light images and corresponding normal-light images. Moreover, many sample pairs are artificially generated, resulting in unnatural enhanced results.

The most popular category of low-light image enhancement methods is based on the Retinex model. Retinex theory models the image depending on the illumination and reflectance components [24]:  $I = R \circ L$ , where  $I$  is the observed image,  $R$  represents the reflectance component of the image, and  $L$  represents the illumination component. The operator  $\circ$  denotes the pixel-wise multiplication. Retinex is a chaining process. The estimation of illumination directly affects that of reflectance. It is essential to estimate the illumination precisely. Variational Retinex-based methods regard the Retinex decomposition as an optimization problem by adding constraint terms on the reflectance and illumination components [27, 36, 46]. Nevertheless, these methods have high computational cost and cannot enhance the contrast well. Most of the existing Retinex-based methods regard  $R$  as the enhanced image [22, 40] and utilize the logarithmic transformation to reduce the computational complexity [7]. In the multi-scale Retinex (MSR) [21], the reflectance is the enhanced image, which is obtained by multi-scale Gaussian filter. However, MSR cannot preserve edge information or enhance the contrast to satisfy human visual system (HVS). In this paper, just-noticeable-difference (JND) model [20] is adopted as an adaptive adjustment for contrast enhancement, and the contrast is combined with JND to decide the variance of the multi-scale adaptive filter, which can make up for the drawback of MSR. In addition, the proposed method generates enhanced results by fusing illumination and reflectance components to preserve the naturalness.

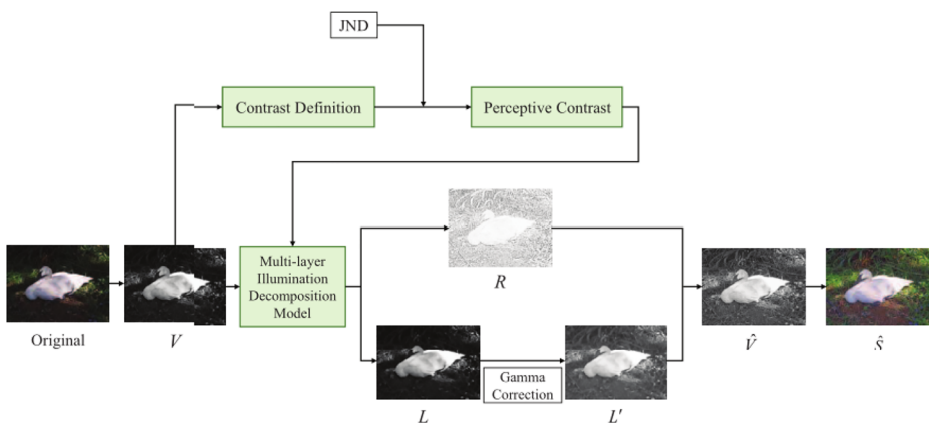
Based on an assumption that details may exist in multi-layer spatial frequency bands, multi-layer model has been adopted in low-light image enhancement [14, 43]. However, these methods cause artifacts, ignore the enhancement of contrast, and cannot satisfy human visual perception. Only a few methods utilize the multi-layer decomposition, and how to apply human visual perception to Retinex multi-layer decomposition is also an open problem.

In this paper, we perform perceptive low-light image enhancement via multi-layer illumination decomposition model. The core concept is to make the illumination and reflectance components region aware. To preserve the hue and saturation, Retinex decomposition is only applied to the value (V) channel of the input image in the HSV color space. First, the contrast is defined from local and global perspectives. At the same time, the effect of noise is integrated into the contrast. Second, the JND model and the defined contrast are combined to generate the perceptive contrast, which can divide the contrast into high and low contrasts adaptively. Then, the multi-scale adaptive filter is proposed, which can enhance high and low contrasts, respectively. Finally, a novel multi-layer illumination decomposition model is presented to extract accurate illumination information, which can also effectively obtain the reflectance component. The final enhanced result is obtained by combining the decomposed multi-layer reflectance and adjusted illumination components. Figure 1 shows the general framework of the proposed method.

- **Contrast Definition:** In Section 3.1, the contrast is defined by considering the noise as well as the global and local characteristics, and it contributes to enhancing the contrast of dark areas while maintaining the contrast of bright areas.
- **Perceptive Contrast:** In Section 3.2, the perceptive contrast is introduced by combining the JND model with the defined contrast. Perceptive contrast can divide the contrast into high and low parts, which require different enhancement.
- **Multi-layer Illumination Decomposition Model:** In Section 3.3, the multi-layer illumination decomposition model is proposed, which can obtain the illumination accurately and preserve the naturalness by iteratively using the multi-scale adaptive filter. The perceptive contrast is utilized to decide the variance of the filter. The decomposed multi-layer spatial frequency bands are fused to obtain the final enhanced result.

Extensive experimental results demonstrate the effectiveness of the proposed method in terms of objective and subjective assessments. The proposed method can improve the contrast and brightness, as well as preserve the naturalness.

The rest of this paper is organized as follows: Section 2 gives a brief survey of related work. The details of the proposed method are provided in Section 3. Section 4 presents the experimental results and compares the proposed method with the state-of-the-arts both in subjective and objective assessments. In addition, parameter study, computational time and



**Fig. 1** The general framework of the proposed method

limitations are also presented in this section. Finally, the work is concluded in Section 5. For ease of reading, the mathematical notations and acronyms used in this paper are listed in Table 1.

## 2 Related work

### 2.1 Human visual perception model

The photoreceptors on the retina of human eyes perceive the light and act as the sensors for HVS [13]. There are two types of photoreceptors, namely, rods and cones. Rods are sensitive to light and responsible for the visual perception in low light conditions (scotopic vision).

**Table 1** Glossary of key acronyms and notations

Symbol	Description
$I$	the observed image
$R$	the reflectance component
$L$	the illumination component
$V$	V channel of the observed image
HVS	human visual system
$JND$	just-noticeable-difference model
$t$	the luminance value of the input image
$g\sigma$	Gaussian kernel with the standard deviation $\sigma$
$c_l(x, y)$	local contrast of the target pixel located at $(x,y)$
$C_g$	global contrast
$c_n(x, y)$	the effect of noise in the pixel located at $(x,y)$
$\hat{C}$	contrast
$\Omega$	the area with a certain distance from the center $(x,y)$
$\tilde{\sigma}_1, \tilde{\sigma}_2$	the parameters used to control the scale
$M$	the number of pixels in $\Omega$
$\alpha, \beta$	geometric proximity and color difference
$s$	scale
$l$	the image to be filtered
$B_{s,l}(x, y)$	the filtered result
$D_i(x, y)$	the detail component in the $i_{th}$ layer
$R_i(x, y)$	the reflectance in the $i_{th}$ layer
$L_i(x, y)$	the illumination in the $i_{th}$ layer
$L_i^\gamma(x, y)$	the adjusted illumination in the $i_{th}$ layer
$w$	the number of layers
$\hat{V}(x, y)$	the enhanced V channel image
$\beta_i$	the weight of $i_{th}$ illumination layer
$\eta_i$	the weight of $i_{th}$ reflectance layer
BTMQI	blind tone-mapped quality index
NIQE	natural image quality evaluator blind image quality assessment
NIQMC	no-reference image quality metric for contrast distortion

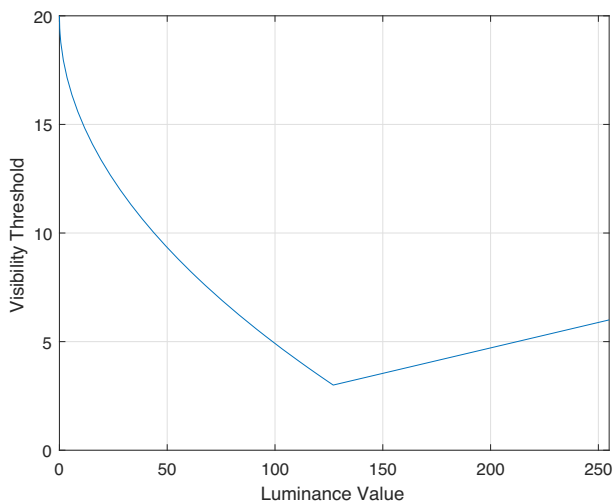
There are three types of cones, which are red, green and blue cones. These cones are used for color vision in bright light conditions (photopic vision), and they are less sensitive to light than rods. The neurons can only transfer signals within a narrow range, while the range that human eyes can adapt to is much wider. HVS has been studied extensively. Karen et al. [33] proposed a multi-histogram equalization method using HVS to segment the image into three regions: Devries-Rose, Weber and saturation regions. This method analyzes the relationship between the background intensities and the contrast threshold. To balance the relationship between HVS and luminance, Jayant [20] proposed a JND model based on the relationship between the actual luminance and the human visual perceptible brightness. JND model can be employed to obtain the smallest brightness difference that HVS can distinguish [51]. If the difference between the actual luminance of an image and the human visual perceptible luminance is smaller than the critical JND, these two luminance values are almost identical for human eyes, *i.e.*, human eyes cannot perceive the difference in this case. The visibility threshold of JND model can be expressed as follows [3]:

$$JND(t) = \begin{cases} 17 \times \left(1 - \sqrt{\frac{t}{127}}\right) + 3 & t \leq 127 \\ \frac{3}{128} \times (t - 127) + 3 & \text{otherwise} \end{cases} \quad (1)$$

where  $t$  is the luminance value of the input image and its variation range is  $[0, 255]$ . As shown in Fig. 2, HVS is more sensitive to bright regions than to dark regions. JND model plays an important role in many fields, such as perceptual image/video compression and image quality measurement.

## 2.2 Low-light image enhancement methods

Histogram-based methods enhance low-light images by modifying the distributions of their histograms. Lim et al. [29] adopted an extended 2D-histogram-based image enhancement scheme to improve the contrast of structure image. Cheng et al. [2] utilized a mapping function to determine the  $\gamma$  value in Gamma correction based on the characteristics of the input histogram.



**Fig. 2** The relationship between luminance value and visibility threshold

Dehazing-based methods obtain the results of low-light image enhancement by inverting the dehazed unrealistic images again. Dong et al. [6] dealt with the inverted video frames by applying an optimized image dehazing algorithm, which utilizes the temporal correlations between subsequent video frames for the estimation of parameters. Li et al. [26] followed this technique to segment the observed image into super pixels and then denoise the segments adaptively using the Block-Matching 3D method (BM3D) [4]. Feng et al. [10] applied a dark channel prior K-means [37] classification method to estimate the transmission and atmospheric light.

Based on the learning principle, many low-light image enhancement networks have been studied. Lore et al. [30] presented a low-light net to enhance the contrast, identify the signal features and brighten images adaptively by using deep autoencoders. Zhang et al. [54] proposed a novel network for kindling the darkness (KinD-Net) to achieve the decomposition of Retinex model. By regarding the low-light image enhancement process as a residual learning problem, Wang et al. [41] proposed a deep lightening network (DLN-Net), which contains several lightening back-projection blocks to learn the residual for normal-light estimation. Different from these end-to-end mapping learning methods, Guo et al. [17] presented a novel image-to-deep curve estimation network (DCE-Net), which outputs high-order curves to adjust the dynamic range and preserve the contrast of neighboring pixels.

Retinex-based methods depend on the illumination and reflectance components. Wang et al. [44] split the observed image into reflectance and illumination components by a bright-pass filter, which can preserve the naturalness and enhance details. Following the bright-pass filter technique, Fu et al. [12] adjusted the illumination map and then fused the derived inputs with the corresponding weights in a multi-scale fusion. The method proposed in [40] is based on a piecewise stretch of the extracted brightness component to improve the visual quality of the image. Wang et al. [43] designed a novel naturalness preserved image enhancement method by using a prior multi-layer lightness statistics (NPIE-MLLS).

The variational Retinex method was initially introduced in [23]. However, since the logarithmic transformation in this method suppresses the variation of gradient magnitude in bright regions, the estimated reflectance component may cause the loss of some details. To overcome the disadvantage appeared in the log-transformed domain, Fu et al. [11] proposed a maximum posteriori method employed in the linear domain to achieve the simultaneous illumination and reflectance estimation (PMSIRE). But the logarithmic transformation can simulate human vision perception of light intensity, so Fu et al. [49] added the reasonable weights into the constraint terms of illumination and reflectance to offset the weakness of logarithmic transformation. In these methods, noise degrades the enhanced results and causes observable halo artifacts. Guo et al. [18] achieved low-light image enhancement via illumination map estimation (LIME), where BM3D is adopted to denoise after adding a structure-aware prior into the initial illumination map. Li et al. [27] proposed a structure-revealing low-light image enhancement method by a robust Retinex model (SRLIE), which regards the noise component as a part of the Retinex model. The noise component can be obtained by solving the variational Retinex objective function. Xu et al. [46] estimated the illumination and reflectance components using a structure- and texture-aware model. Ren et al. [36] injected the low-rank regularized prior into the Retinex decomposition model (LR3M) to suppress the noise contained in the reflectance map. Hao et al. [19] estimated the illumination and reflectance components in a semi-decoupled decomposition way (SDD), and denoised during the estimation of the reflectance.

### 3 Proposed method

This section presents the proposed method in detail. First, the definition of contrast is provided. Then, the perceptive contrast is stated by combining the JND model with the defined contrast to satisfy the human visual perception. Finally, the multi-layer illumination decomposition model is introduced.

#### 3.1 Contrast definition

In this paper, the contrast is defined from both local and global aspects. Furthermore, the effect of noise is considered. The definition of contrast is introduced in detail below.

First, the contrast defined in [8] is adopted as the local contrast in this paper, which is estimated as a standard deviation computed within a Gaussian window, and its expression is as follows:

$$c_l(x, y) = \sqrt{(g_\sigma * V^2)(x, y) - (g_\sigma * V)^2(x, y)} \quad (2)$$

where  $c_l(x, y)$  is the local contrast of the target pixel located at  $(x, y)$ .  $*$  is the convolution operator,  $V$  is the V channel of the observed image in the HSV color space, and  $g_\sigma$  is a Gaussian kernel with the standard deviation  $\sigma$  (set to be 3). The matrix  $C_l$  is adopted as the local contrast of the entire image.

Second, the average and variance values are utilized to represent the global contrast, since they reflect the global information. The global contrast is expressed mathematically as follows:

$$C_g = V_{var} - \bar{V}^2 \quad (3)$$

where  $V_{var}$  and  $\bar{V}$  are the variance and average values of V channel in the observed image, respectively.  $C_g$  indicates the global contrast.

Third, the effect of noise is considered. Since HVS perceives visual information based on the grey-level differences of adjacent pixels instead of their absolute grey levels [5, 9], the relationship between adjacent pixels is analyzed and the target grey-level difference is computed to avoid the noise amplification. The effect of noise is introduced as follows:

$$c_n(x, y) = \frac{1}{M} \sum_{(m,n) \in \Omega} \exp\left(-\frac{(V(m, n) - V(x, y))^2}{2\tilde{\sigma}_1^2} - \frac{(m-x)^2 + (n-y)^2}{2\tilde{\sigma}_2^2}\right) \quad (4)$$

where  $\Omega$  represents the area with a certain distance (empirically set to 3) from the center  $(x, y)$ ,  $M$  is the number of pixels in  $\Omega$ , and  $(m, n)$  is the location of pixels in  $\Omega$ .  $\tilde{\sigma}_1$  and  $\tilde{\sigma}_2$  are empirically set to be 2 to control the scale. The first term represents the pixel-wise difference between the neighbouring area and the center, and the second term is utilized to constrain the distance between  $(m, n)$  and  $(x, y)$ . These two terms stand for the grey similarity and positional proximity, respectively.  $c_n(x, y)$  represents the effect of noise in the pixel located at  $(x, y)$ . When the pixel is close to  $(x, y)$  and its grey value is close to  $V(x, y)$ , the corresponding  $c_n(x, y)$  has a high value. Conversely,  $c_n(x, y)$  has a low value.  $C_n$  is denoted as the effect of noise on the whole image.

Equation (4) illustrates that the nearer the pixels are, the higher sensibility does HVS present to the grey-level difference. If the second term is fixed,  $c_n(x, y)$  completely depends on the first term.  $c_n(x, y)$  satisfies the principle that its value is lower when the pixel contains more noise than other pixels. That is to say, when the pixel is degraded due to excessive noise, the noise poses less effect on pixels which need enhancement. For example, as shown in Fig. 3, the pixel  $p_1$  is within a flat region, which is not likely to be noisy, so it needs to be assigned with a high  $c_n$  value. In contrast,  $q_1$  is located in a noisy region and it needs to

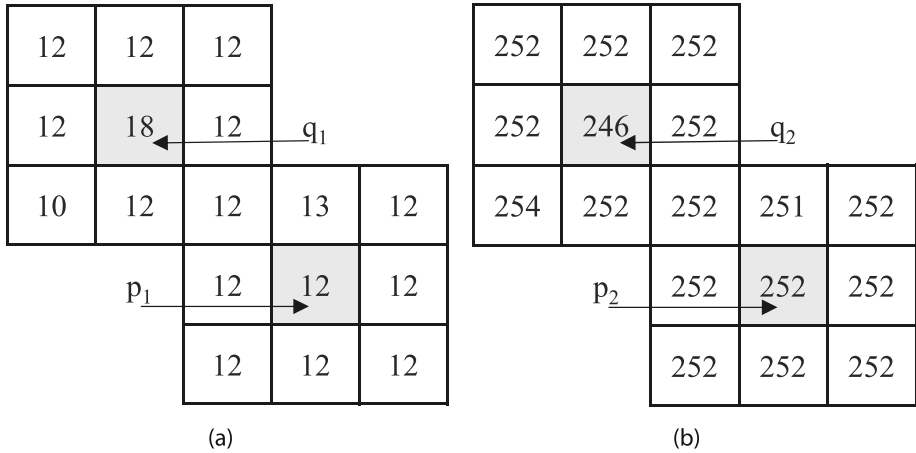


Fig. 3 An example of the effect of noise  $c_n$

be assigned with a low  $c_n$  value. When the differences between the pixels of neighbouring areas and  $V(x, y)$  are less than 2, *i.e.*,  $|V(m, n) - V(x, y)| \leq 2$ , these pixels are regarded as at the similar grey levels.

Finally, the contrast  $\hat{C}$  of the target image is obtained by multiplying the matrix  $C_n$  with  $C_g$  and  $C_l$ , as follows:

$$\hat{C} = C_g \cdot C_l \circ C_n \tag{5}$$

As shown in Fig. 3,  $p_2$  and  $q_2$  have the same  $c_n$  values as  $p_1$  and  $q_1$ , respectively. However, the contrast  $\hat{C}_{p_1q_1}$  is much larger than  $\hat{C}_{p_2q_2}$ . This demonstrates that the defined contrast can not only enhance the contrast of dark areas, but also maintain the contrast of bright areas.

### 3.2 Perceptive contrast

MSR performs the convolution between the Gaussian smoothing function and the original image to obtain the illumination layer. But MSR may cause halo artifacts along strong edges due to the assumption of smooth illumination. Therefore, in this paper, an adaptive filter is designed to reduce halo artifacts by adjusting the shape of the filter for high-contrast and low-contrast areas. How to divide the defined contrast into high and low contrasts is the key.

As illustrated in Section 2.1, the JND model satisfies the requirements of the human visual perception system, and it can also show the relationship between the visibility threshold and the luminance. In this paper, the defined contrast and the JND model are combined to obtain the threshold used in the contrast division, which is called perceptive contrast.

When the contrast is higher, HVS perceives the changes more sensitively in the case of the fixed luminance, *i.e.*, the visibility threshold of HVS is lower. Therefore, it is reasonable to assume that the relationship between the contrast and the visibility threshold is inversely proportional, expressed as follows:

$$\hat{C} \propto \frac{k}{JND} \tag{6}$$



where  $k$  is the proportionality coefficient.  $JND$  can be represented as  $JND = \frac{k}{\hat{C}}$ . Putting this derivation into (1) can make the following conclusion:

$$\frac{k}{\hat{C}} = \begin{cases} 17 \times (1 - \sqrt{\frac{t}{127}}) + 3 & t \leq 127 \\ \frac{3}{128} \times (t - 127) + 3 & \textit{otherwise} \end{cases} \tag{7}$$

Equation (7) is a piecewise function. There are two situations that need to be discussed:

1) **When  $t \leq 127$ :**

The mathematical expression of  $t$  is derived as follows:

$$t = (1 - \frac{\frac{k}{\hat{C}} - 3}{17})^2 \times 127 \tag{8}$$

When  $t \leq 127$ , it means that the target image is in low light conditions. According to (8), a constraint of  $t$  can be made:  $(1 - \frac{\frac{k}{\hat{C}} - 3}{17})^2 \leq 1$ . Then,  $0 < 1 - \frac{\frac{k}{\hat{C}} - 3}{17} \leq 1$  is workable, which can further restrict the range of  $\hat{C}$ . Based on the above analysis, the range of  $\hat{C}$  is  $\frac{k}{20} \leq \hat{C} < \frac{k}{3}$ , under the condition of  $t \leq 127$ .

2) **When  $t$  is other value:**

The formula of  $t$  is expressed as follows:

$$t = (\frac{k}{\hat{C}} - 3) \times \frac{128}{3} + 127 \tag{9}$$

Since  $t$  is the luminance value of the input image within the range of  $[0, 255]$ , other values refer to the range of  $[128, 255]$  for 8 bit images. Similar to the above analysis when  $t \leq 127$ , the range of  $\hat{C}$  is  $[\frac{k}{6}, \frac{k}{3})$ , under the condition that  $t$  is other value.

In summary, the perceptive contrast is a piecewise function and  $\frac{k}{6}$  is the demarcation value to divide the contrast into high or low ones. The pixels whose  $\hat{C}$  values are in the range  $[\frac{k}{20}, \frac{k}{6})$  have low contrast and need more enhancement. On the contrary, the pixels whose  $\hat{C}$  values are in the range  $[\frac{k}{6}, \frac{k}{3})$  have relatively high contrast and need less enhancement.

### 3.3 Multi-layer illumination decomposition model

#### 3.3.1 Illumination decomposition by multi-scale adaptive filter

Different from the Gaussian filter, the bilateral filter not only uses the geometric proximity, but also takes the color difference between pixels into account. This is why the bilateral filter can effectively remove noise and preserve the edge information of the image. In addition, the bilateral filter is non-iterative and regional. Therefore, the bilateral filter is adopted in this paper instead of the Gaussian filter used in MSR to estimate the illumination component, which can reduce the effect of noise and avoid halo artifacts.

Inspired by the above, a multi-scale adaptive filter is proposed to estimate the illumination. The estimated illumination component is expressed as follows:

$$B_{s,l}(x, y) = \frac{\sum_{(p,q) \in \tilde{\Omega}} \alpha\beta l(p, q)}{\sum_{(p,q) \in \tilde{\Omega}} \alpha\beta} \tag{10}$$

where  $\tilde{\Omega}$  represents the areas of multi-scale distance from the center  $(x, y)$ , and  $(p, q)$  represents the location of pixels within  $\tilde{\Omega}$ .  $s$  stands for the characteristic of multi-scale,

and it takes on the values of 1, 2 and 4, representing the small, middle and large scales, respectively.  $B_{s,l}(x, y)$  is the result of filtering the image  $l$  by  $s$  scale.  $\alpha$  and  $\beta$  are the geometric proximity and color difference between pixels, expressed as follows:

$$\alpha = e^{-\frac{(p-x)^2+(q-y)^2}{2\sigma_n^2}} \tag{11}$$

$$\beta = e^{-\frac{|l(x,y)-l(p,q)|^2}{2\sigma_l^2}} \tag{12}$$

where  $\sigma_n$  and  $\sigma_l$  represent the geometric variance and color variance, respectively. In this paper,  $\sigma_n$  and  $\sigma_l$  are set to be equal, and their values are constrained by the perceptive contrast so that halo artifacts can be adaptively removed in the estimation of illumination.  $\sigma_n$  and  $\sigma_l$  are defined as follows:

$$\sigma_n = \sigma_l = \begin{cases} \sigma_0 & \frac{k}{20} \leq \hat{c}(x, y) < \frac{k}{6} \\ \sigma_1 & \text{otherwise} \end{cases} \tag{13}$$

where  $\hat{c}(x, y)$  represents the contrast of the pixel located at  $(x, y)$ . From (13), if the contrast is less than  $\frac{k}{6}$  and more than  $\frac{k}{20}$ ,  $\sigma_0$  is used to reduce the noise in the smooth areas; otherwise,  $\sigma_1$  is used to maintain the edge information.

### 3.3.2 Multi-layer model

Multi-layer model is based on an essential assumption that details may exist in the multi-layer spatial frequency bands. Therefore, the multi-scale adaptive filter is used to iteratively decompose the input image into multiple layers until the illumination map is uniform.

Figure 4 shows the overview of the multi-layer model. Each component  $D_i(x, y)$  is further decomposed into a low-frequency component  $L_{i+1}(x, y)$  and a detail component  $D_{i+1}(x, y)$ .  $L_{i+1}(x, y)$  is computed by filtering  $D_i(x, y)$  using the multi-scale adaptive filter, as follows:

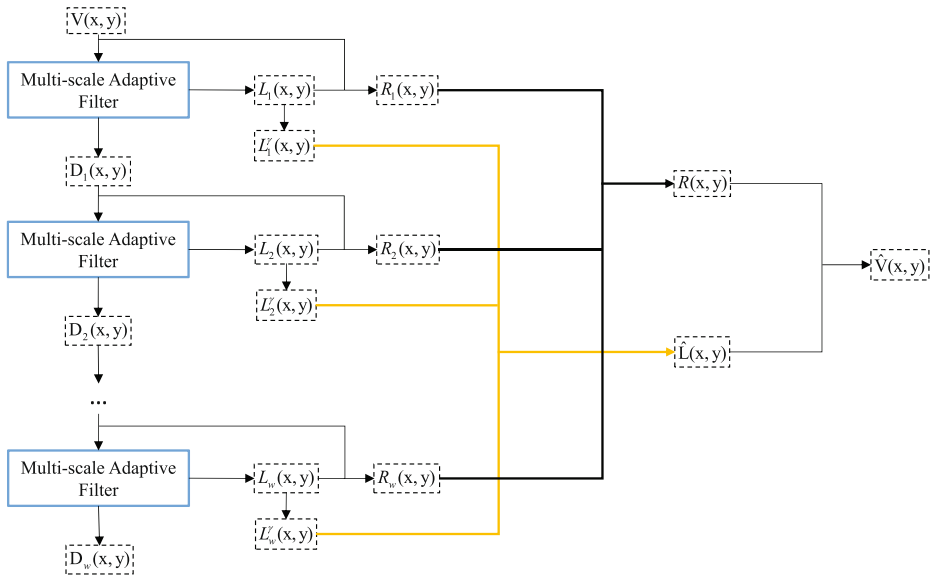
$$L_{i+1}(x, y) = \sum_s \xi_s B_{s,D_i}(x, y) \tag{14}$$

where  $\xi_s$  is the weight coefficient of each scale.  $\xi_1, \xi_2$  and  $\xi_4$  are set to be 0.33, 0.34 and 0.33, respectively. Each detail component  $D_{i+1}(x, y)$  is obtained by  $D_i(x, y) - L_{i+1}$ . At the same time, each reflectance component  $R_{i+1}(x, y)$  is obtained by  $R_{i+1}(x, y) = D_i(x, y)/L_{i+1}(x, y)$ . The brightness range of the low-frequency component shrinks as  $i$  increases. If the layer number is large enough, the brightness of the last layer will be completely uniform. Each  $R_i(x, y)$  covers an unique frequency band of the original image, which can be integrated into the final enhanced image to preserve details.

In the proposed multi-layer model, the image  $V(x, y)$  is decomposed into multiple components, including multiple illumination and reflectance maps. After obtaining the reflectance and illumination components, the next task is to adjust the illumination to improve the visibility. In this paper, Gamma correction is applied. The enhanced V channel image  $\hat{V}$  is generated by:

$$\hat{V}(x, y) = \left(\sum_{i=1}^w \beta_i L_i^\gamma(x, y)\right) \cdot \left(\sum_{i=1}^w \eta_i R_i(x, y)\right) \tag{15}$$

where  $w$  represents the number of layers.  $\gamma$  is empirically set as 2.2. When  $i = 1$ ,  $L_1(x, y)$  is obtained by directly decomposing the original V channel image  $V(x, y)$ . With the increase of the layer number, the illumination and the corresponding reflectance components tend to be closer to the actual illumination and reflectance, respectively. Therefore, their weights ( $\beta_i$



**Fig. 4** The overview of multi-layer model.  $D_i(x, y)$  indicates the detail component in the  $i_{th}$  layer;  $L_i(x, y)$  and  $R_i(x, y)$  represent the illumination and reflectance components in the  $i_{th}$  layer, respectively;  $L_i^\gamma(x, y)$  indicates the adjusted illumination;  $R(x, y)$  and  $\hat{L}(x, y)$  stand for the integrated reflectance and illumination, respectively;  $\hat{V}(x, y)$  is the enhanced V channel image

and  $\eta_i$ ) should increase as  $i$  increases, which are set to be  $\frac{2i-1}{w^2}$ . The whole decomposition will be stopped only if the difference between  $L_i(x, y)$  and  $L_{i+1}(x, y)$  (or between  $R_i(x, y)$  and  $R_{i+1}(x, y)$ ) is smaller than a threshold in practice, *i.e.*,  $\| L_{i+1}(x, y) - L_i(x, y) \| / \| L_i(x, y) \| < 5 \times 10^{-3}$  (or  $\| R_{i+1}(x, y) - R_i(x, y) \| / \| R_i(x, y) \| < 5 \times 10^{-3}$ ), or if the maximal number of layers is reached. The illumination in (15) does not need to be normalized before Gamma correction since the input V channel image  $V$  is converted to normalization. Finally, the enhanced HSV image is transformed to RGB color space, generating the final result  $\hat{S}$ .

### 4 Experimental results and analysis

First, the experiment settings are presented in this section. Second, the proposed method is evaluated by comparing it with the state-of-the-art low-light image enhancement algorithms to demonstrate the efficiency and effectiveness of our method. Third, the parameters used in the proposed method are analyzed. Fourth, the computational time of the proposed method is discussed. Finally, the limitations and future directions of our work are also pointed out.

#### 4.1 Experiment settings

To comprehensively evaluate the proposed method, we test the proposed method on images with different illumination conditions. The test images are obtained from NPE [44], LIME [18], MEF [31], MF [12] and DICM [25] datasets. All images from these datasets have low contrast in local areas but serious illumination variation in global space.

If not specifically stated, the parameter  $k$  in the expression (6) is 1, and the parameters  $\sigma_0$  and  $\sigma_1$  in (13) are 3.3 and 0.5, respectively. In addition, the parameter  $w$  in (15) is set to be 10. These parameters are analyzed in Section 4.3.

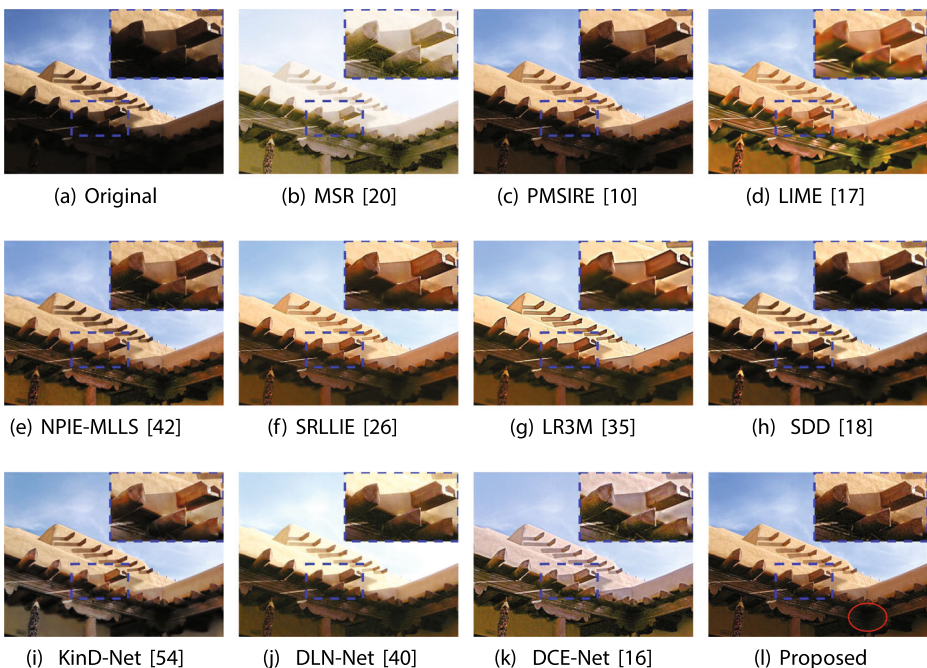
## 4.2 Quality assessment

The proposed method is compared with several state-of-the-art methods, including seven traditional algorithms (MSR [21], PMSIRE [11], LIME [18], NPIE-MLLS [43], SRL-LIE [27], LR3M [36] and SDD [19]) and three learning-based networks (KinD-Net [54], DLN-Net [41] and DCE-Net [17]). Experiments of traditional methods are performed in MATLAB R2016b on a computer running Windows 10 with an Intel (R) Core (TM) i5-3470 CPU and 8.00GB of RAM. The results of these methods are generated using the codes downloaded from the authors' websites, and the settings of parameters are the same as the original papers. Learning-based methods are trained with datasets in [17] and performed on the Nvidia RTX 2080Ti GPU with 11G memory.

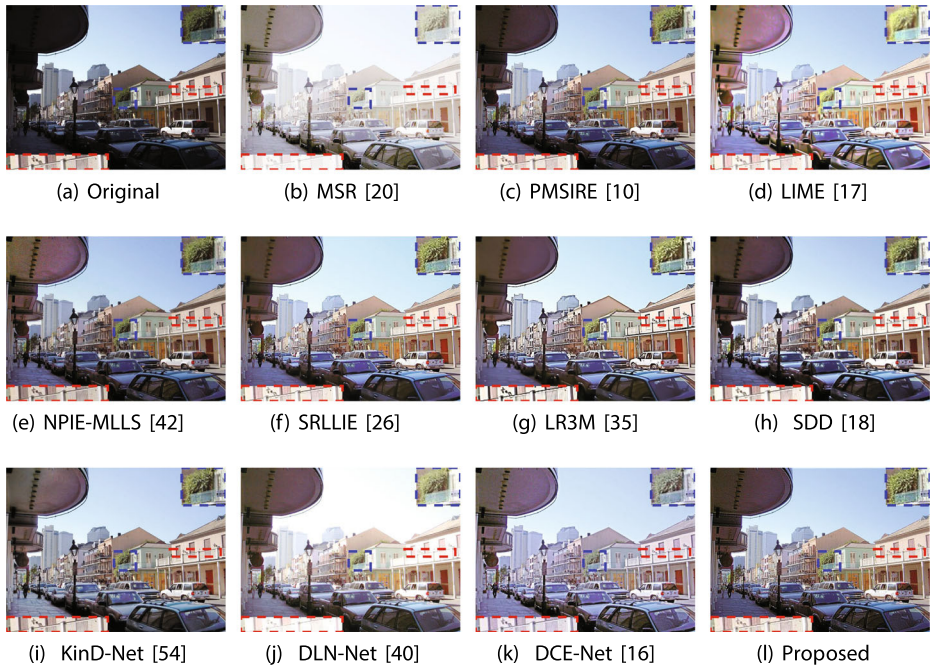
### 4.2.1 Subjective assessment

Due to the space limitation, this section only presents some representative results as shown in Figs. 5, 6, 7, 8, 9 and 10.

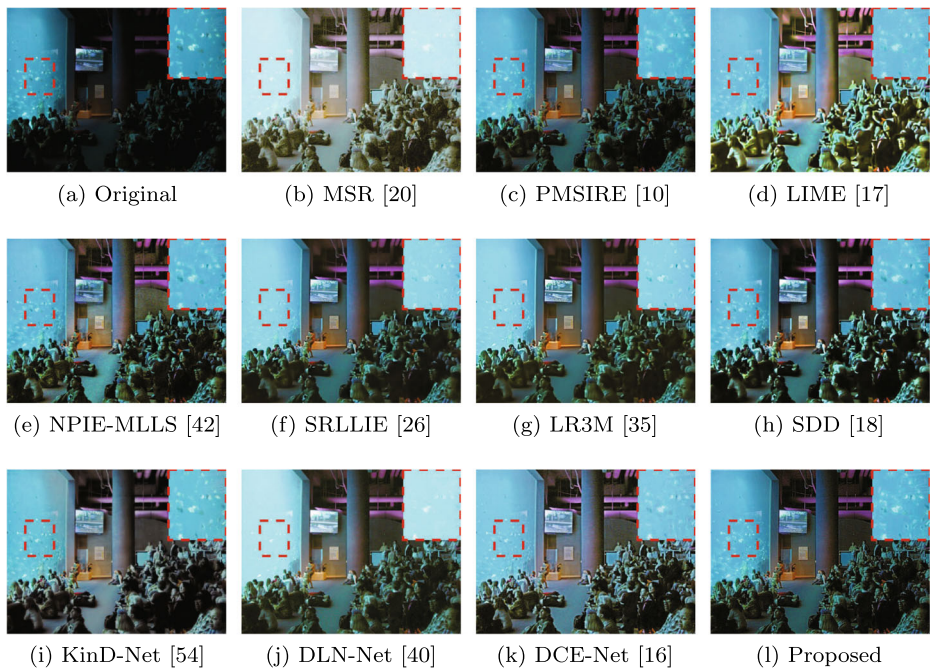
Since MSR simply takes the reflectance as the final result, the enhanced images shown in Figs. 5(b)–10(b) do not preserve the naturalness. For example, the color of the sky is close to white in Fig. 6(b). In addition, since the Gaussian filter cannot maintain the contrast of edges, the enhanced results by MSR tend to lose some details, such as the sparse texture



**Fig. 5** Enhanced results of *Eave* with different methods



**Fig. 6** Enhanced results of *Street* with different methods



**Fig. 7** Enhanced results of *Aquarium* with different methods

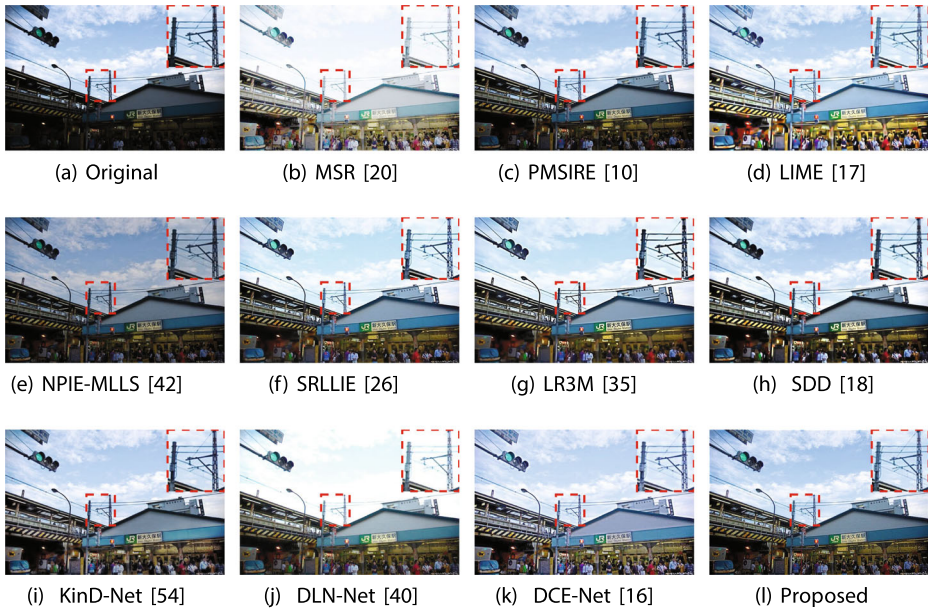


Fig. 8 Enhanced results of *Station* with different methods



Fig. 9 Enhanced results of *Park* with different methods



**Fig. 10** Enhanced results of *Rock* with different methods

of rocks in Fig. 10(b). Compared with MSR, the proposed algorithm adopts the perceptual contrast to enhance the contrast adaptively, which can preserve the naturalness, maintain detail information and improve the contrast well.

PMSIRE incorporates the Bayes' theorem into Retinex theory to formulate a probabilistic model, which can simultaneously estimate the reflectance and illumination components in the linear domain. Figures 5(c)–10(c) show the results of PMSIRE. The overall brightness of the enhanced images is not well improved, even though the enhanced images show some impressive results. For example, in Fig. 5(c), the area under the eave is still buried in darkness. Moreover, PMSIRE cannot enhance the contrast, as shown the smooth structure of aquarium in Fig. 7(c). Comparatively, the proposed algorithm achieves a good tradeoff among detail, contrast, brightness and naturalness.

LIME is an effective method to adjust the brightness of low-light images, which adds a structure-aware prior into the initial illumination map. As shown in Figs. 5(d)–10(d), the enhanced results have good brightness information. Nevertheless, owing to the lack of constraint on the reflectance, this method can easily over enhance the high-intensity regions, such as the over-exposed areas of glass in *Aquarium*. Furthermore, the enhanced images exist color distortion, such as the color of roof in Fig. 6(d) and grass in Fig. 9(d). The color distortion has a negative impact on human visual perception, so that the enhanced image loses naturalness. LIME utilizes BM3D method to reduce the effect of noise. However, denoising may cause the results fuzzy, as shown by the profile of pedestrians in *Station*. In contrast, the proposed method generates more natural results while also considering the influence of noise.

NPIE-MLLS attempts to preserve the naturalness of images. It is effective in improving the local contrast in dark areas and keeping the lightness-order of the image. As shown in

Figs. 5(e)–10(e), most of results are satisfactory. However, the brightness is not improved, such as the sky in Fig. 8(e). NPIE-MLLS cannot keep spatial consistency with the original image since there is no consideration of human visual perception. In Fig. 8(a), the values of pixels in the sky area are higher than that in other areas. From the enhanced result in Fig. 8(e), although the brightness of the areas other than the sky is increased, the brightness of the sky is reduced. In addition, there are some artifacts, such as the black areas in the balcony of the green building in Fig. 6(e). Unlike NPIE-MLLS, the proposed method can successfully preserve the spatial consistency with the original image and avoid artifacts by using the human visual characteristic.

SRLIE presents a robust Retinex model by considering the effect of noise, and utilizes the gradient of reflectance to reveal structural details. SRLIE can estimate the reflectance, illumination and noise components effectively. From Figs. 5(f)–10(f), it is obvious that SRLIE can obtain satisfactory results in most cases. Nevertheless, the contrast is not enhanced effectively. For example, the contrast of cars in Fig. 6(f) is low, and the frame of eave in Fig. 5(f) is over-smooth and blurry. This is because SRLIE only applies the local characteristic into Retinex variational model, while overlooking the contrast enhancement. Furthermore, the hue of enhanced images by SRLIE is not natural. As shown in Fig. 8(f), color distortion exists in the sky. Compared with SRLIE, the enhanced images by the proposed method have higher contrast and more natural brightness.

LR3M injects low-rank prior to suppress the noise in the reflectance component based on the Retinex variational model. Figures 5(g)–10(g) show the results of LR3M. The estimated illumination information by LR3M is too smooth, which causes the loss of detail information in the enhanced results. For example, in Fig. 5(g), the texture of eave is not distinguishable. The structural information is not clear either, such as the outline of cars in Fig. 6(g). Due to the lack of color constraint in LR3M, slight color distortion exists in the enhanced results. For instance, the sky in Fig. 8(g) tends to be light blue. Different from LR3M, our method maintains color consistency and enhances the contrast by applying the perceptive contrast. Furthermore, the naturalness is also well preserved in the enhanced results by the proposed method.

SDD estimates illumination and reflectance components by a semi-decoupled decomposition model, which obtains the illumination by an edge-preserving image filter and suppresses the noise during the reflectance estimation. The enhanced results of SDD are shown in Figs. 5(h)–10(h), which are natural but lose some details. For example, the texture of the sorghum in Fig. 5(h) is indistinct. The outline of cars shown in Fig. 6(h) is blurred, leading to the low contrast of the entire image. This is because SDD over smooths the illumination information. Moreover, SDD does not balance the relationship between brightness and contrast, which causes unreasonable exposure. As shown in Figs. 5(h) and 9(h), the sky areas are over-exposed. Compared with SDD, the proposed method enhances the brightness naturally and improves the contrast reasonably by utilizing the characteristic of HVS.

KinD-Net is a deep network designed to decompose the Retinex model, which is trained with paired samples captured under different light conditions. Figures 5(i)–10(i) are the enhanced images by KinD-Net. The brightness of these enhanced results are satisfactory, but these results exist color distortion since the illumination adjustment net of KinD-Net lacks the natural samples. As shown in Fig. 6(i), the floor is off-white. The water in the aquarium shown in Fig. 7(i) is bluish green, which is not consistent with the hue of the original image. Furthermore, KinD-Net does not improve detail information as the restoration net ignores the contrast, such as the low contrast of Fig. 8(i). It is obvious that the sky area in Fig. 9(i) exists artifacts. Different from KinD-Net, the proposed method can avoid color distortion, reduce artifacts, and maintain detail information.



DLN-Net is an end-to-end structure based on the convolutional neural network, which contains several lightening blocks and feature aggregations to enhance the low-light image from global and local aspects. From Figs. 5(j)–10(j), it is obvious that DLN-Net results in over-enhancement, especially in the sky area. Over-exposure causes the loss of details, such as the blurred balustrade in Fig. 6(j). This is because that samples are synthetic. Moreover, these synthetic images lack the naturalness, leading to the color distortion, such as the drain holes in Fig. 10(j). Unlike DLN-Net, the proposed method can avoid over-exposure, improve detail information, and preserve the naturalness simultaneously.

DCE-Net enhances the low-light image by an image-to-curve network, which is trained without any paired samples. It is novel to produce the high-order curves to adjust the dynamic range of the input image. Nevertheless, since some over-exposed images are used in the training set, the enhanced images are in the gray tone, as shown in Figs. 5(k)–10(k). For example, the sky in Fig. 5(a) is blue, while it becomes gray blue after enhancement by DCE-Net shown in Fig. 5(k). Furthermore, the detail information is not improved well. As shown in Fig. 7(k), many fish in the aquarium are too misty to see. Compared with DCE-Net, the enhanced results by the proposed method show more details and keep the color consistency.

These results demonstrate that the proposed method can improve the brightness, enhance the contrast, preserve the naturalness and reduce halo artifacts.

#### 4.2.2 Objective assessment

Since the ground truth of the enhanced image is unknown for a low-light image, several blind objective assessments are employed to quantify the quality, *i.e.*, the blind tone-mapped quality index (BTMQI) [15], the natural image quality evaluator blind image quality assessment (NIQE) [32], and the no-reference image quality metric for contrast distortion (NIQMC) [16]. These matrices are widely used to measure the visual quality from different aspects. The evaluation of contrast uses NIQMC and BTMQI. Meanwhile, NIQE assesses the naturalness of the enhanced results. Tables 2, 3 and 4 show the objective assessment results on the five test datasets. The best objective value is bolded, and the second best value is underlined.

The first assessment metric is BTMQI, which measures the average intensity, contrast and structure information of the tone-mapped image. A smaller BTMQI value represents

**Table 2** Quantitative measurement results of BTMQI

Methods	NPE [44]	LIME [18]	MEF [31]	MF [12]	DICM [25]	Average
MSR [21]	4.985	3.668	4.568	5.396	5.173	4.758
PMSIRE [11]	3.694	4.061	4.339	4.039	4.472	4.121
LIME [18]	4.609	3.181	4.611	4.672	4.714	4.357
NPIE-MLLS [43]	3.311	4.467	3.557	3.263	3.849	<u>3.689</u>
SRLLE [27]	4.115	3.434	3.946	3.763	4.668	3.985
LR3M [36]	3.877	3.750	4.196	3.751	4.759	4.067
SDD [19]	3.728	3.636	4.025	3.783	4.573	3.949
KinD-Net [54]	3.627	3.212	3.759	3.530	4.353	3.696
DLN-Net [41]	4.534	3.239	4.228	4.240	5.173	4.283
DCE-Net [17]	4.070	3.319	4.075	4.201	4.535	4.040
Proposed	3.622	3.227	3.765	3.483	4.343	<b>3.688</b>

**Table 3** Quantitative measurement results of NIQE

Methods	NPE [44]	LIME [18]	MEF [31]	MF [12]	DICM [25]	Average
MSR [21]	3.037	3.819	3.161	3.432	3.670	3.424
PMSIRE [11]	2.892	4.023	3.237	3.319	3.643	<u>3.423</u>
LIME [18]	3.720	4.619	4.383	4.107	4.237	4.213
NPIE-MLLS [43]	2.793	6.582	2.885	3.052	3.423	3.747
SRLLIE [27]	4.199	4.023	4.145	4.009	4.245	4.124
LR3M [36]	3.827	4.429	4.713	4.484	4.551	4.401
SDD [19]	2.988	3.998	3.743	3.548	3.933	3.642
KinD-Net [54]	3.051	4.763	3.823	3.445	4.138	3.844
DLN-Net [41]	3.447	4.165	3.348	3.754	3.636	3.670
DCE-Net [17]	3.102	3.769	3.352	3.440	3.560	3.445
Proposed	3.031	3.807	3.155	3.332	3.641	<b>3.393</b>

better image quality. From Table 2, the average BTMQI value of the proposed method on five datasets is the smallest among all methods for comparison. This indicates that the proposed method can preserve the naturalness of the image in terms of illumination, contrast, structure, color and scene.

The second metric is NIQE, which assesses the image quality by measuring the distance between the model statistics extracted from natural and undistorted images. A lower NIQE represents higher image quality. As shown in Table 3, the proposed method has the lowest average value of NIQE among all algorithms on all test datasets. This means the results of the proposed method are the closest to natural images and have the least distortion.

The last metric is NIQMC, which utilizes local detail and global histogram to assess the quality of images. For NIQMC, a larger value indicates better quality in terms of contrast. As shown in Table 4, LIME has the highest average NIQMC value, which is 0.062 higher than that of the proposed method. However, LIME exits over-exposure, such as the glass of aquarium in Fig. 7(d), and it has poorer performances than our method in other objective assessments.

**Table 4** Quantitative measurement results of NIQMC

Methods	NPE [44]	LIME [18]	MEF [31]	MF [12]	DICM [25]	Average
MSR [21]	4.848	5.083	5.117	4.788	4.857	4.939
PMSIRE [11]	5.110	4.559	4.929	4.770	4.932	4.860
LIME [18]	5.112	5.064	5.233	5.220	5.007	<b>5.127</b>
NPIE-MLLS [43]	5.191	4.056	5.242	5.179	5.118	4.957
SRLLIE [27]	4.994	4.815	5.460	4.832	4.928	5.006
LR3M [36]	5.083	4.751	4.946	4.810	4.980	4.914
SDD [19]	5.066	4.849	5.110	5.017	5.175	5.043
KinD-Net [54]	5.089	4.942	5.167	4.939	5.016	5.031
DLN-Net [41]	5.069	4.980	5.021	5.019	4.909	5.000
DCE-Net [17]	5.061	4.839	5.042	4.917	4.934	4.959
Proposed	5.181	4.985	5.157	5.019	4.983	<u>5.065</u>

Furthermore,  $NNB$  [45] is utilized to test the performance of these methods comprehensively, which integrates BTMQI, NIQE and NIQMC as follows:

$$NNB = 2 \times NIQMC - NIQE - BTMQI \quad (16)$$

For  $NNB$ , a larger value represents better image quality. The test results are shown in Fig. 11. The proposed method has the highest  $NNB$  value, which further demonstrates that our method is the best one among compared algorithms in terms of objective perspective.

In summary, the proposed method generates images with better visual quality both in subjective and objective assessments in most cases. The enhanced images have almost no artifacts and show improvement on the aspects of detail, brightness and naturalness.

### 4.3 Parameter study

Figure 12 shows 12 examples of test images, which are randomly collected from the above datasets and NASA images<sup>1</sup> to demonstrate the reliability of parameters in the proposed method.

#### 4.3.1 Parameter $k$

Since the contrast varies from 0 to 1, the inverse proportional parameter  $k$  in the expression (6) varies from 1 to 3. The enhanced images with different  $k$  values are evaluated by three objective assessment indexes. Figure 13 shows the average objective assessment results of 12 example images shown in Fig. 12. Compared with the case where  $k$  is other value, when  $k$  is 1, the average NIQMC value is the highest, and the average BTMQI and NIQE values are the smallest. When  $k$  increases from 2 to 3, the objective assessments tend to be stable. Therefore, in this paper,  $k$  is set to be 1, which can achieve the enhancement of low-light images well.

#### 4.3.2 Parameters $\sigma_0$ and $\sigma_1$

The impact of parameters  $\sigma_0$  and  $\sigma_1$  in (13) is shown in Fig. 14. The average objective results are obtained from 12 example test images with different  $(\sigma_0, \sigma_1)$  pairs.  $(\sigma_0, \sigma_1)$  takes on the pair-values of (3.3, 0.5), (3.3, 0.05), (33, 0.5) and (33, 0.05), respectively. From Fig. 14, when  $(\sigma_0, \sigma_1)$  is (3.3, 0.5), NIQE and BTMQI have the lowest average values. At the same time, the corresponding NIQMC has the largest average value. When  $(\sigma_0, \sigma_1)$  is (3.3, 0.05), (33, 0.5) and (33, 0.05), the corresponding average NIQE and BTMQI values are large, and the average NIQMC value is small. This is because the order of magnitude of  $\sigma_0$  is much larger than that of  $\sigma_1$ , which makes the bilateral filter play no important role in avoiding halo effects. Therefore,  $\sigma_0$  and  $\sigma_1$  are set to be 3.3 and 0.5 respectively, which can generate satisfactory results in most cases.

#### 4.3.3 Parameter $w$

The effect of the layer number  $w$  in (15) is analyzed using the image of Fig. 12(k). The layer number  $w$  varies from 1 to 15. Figure 15 shows the results of three blind objective assessments with different layer numbers. When the layer number is 10, the corresponding NIQMC has the largest value, while NIQE and BTMQI have the lowest values. When the

<sup>1</sup><https://dragon.larc.nasa.gov/retinex/pao/news>

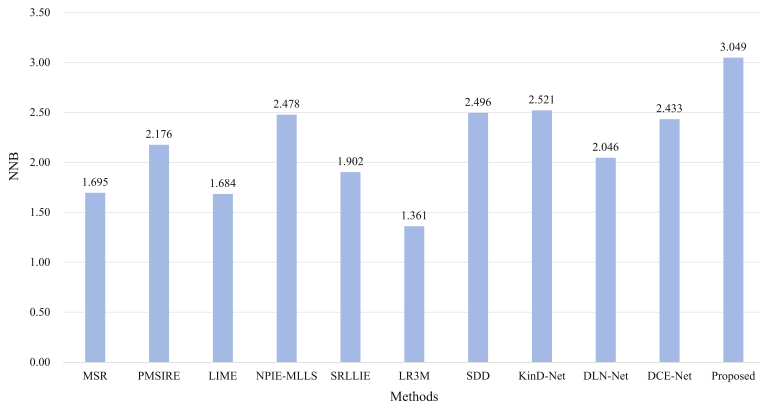


Fig. 11 NNB values of compared methods

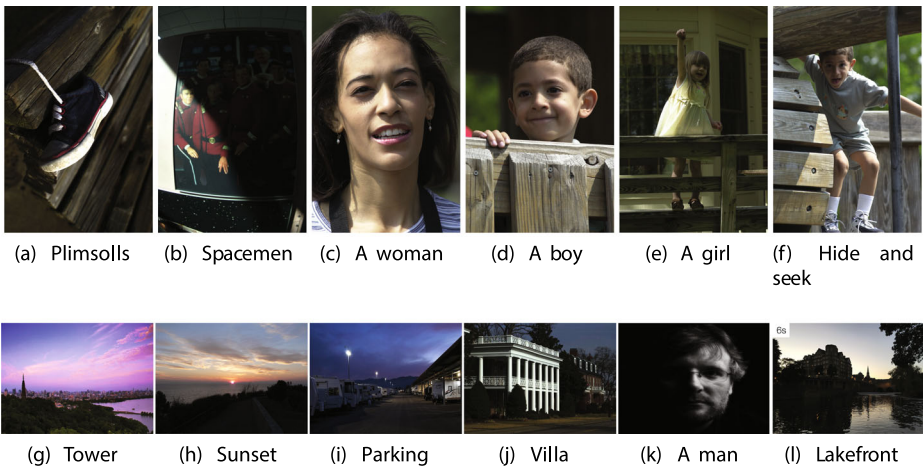


Fig. 12 Examples of test images

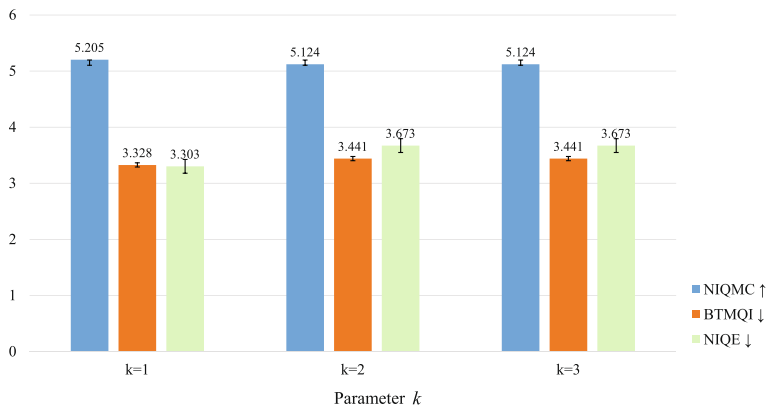


Fig. 13 Average objective assessments with different  $k$

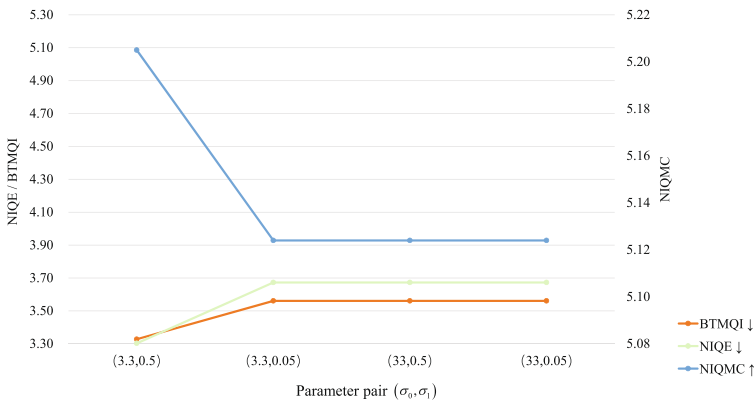


Fig. 14 Average objective assessments with different parameter pairs ( $\sigma_0, \sigma_1$ )

layer number increases from 10 to 15, the objective assessments are stable. This means that the proposed method can achieve the best results when the layer number is 10. Therefore, in this paper, the parameter  $w$  is set to be 10. In addition, Fig. 15 demonstrates that the multi-layer model has an advantage over the single-layer model.

Figure 16 plots the convergence curves for 12 example test images, and gives an intuitive demonstration of the convergence speed of the proposed method. From the curves, the algorithm converges within 10 layers for 12 example test images. Setting the maximum number of layers to be 10 is sufficient to achieve convergence.

### 4.4 Computing efficiency

Since the proposed method belongs to Retinex-based algorithms, the comparison of computing efficiency only performs on seven traditional state-of-the-art methods (MSR [21], PMSIRE [11], LIME [18], NPIE-MLLS [43], SROLLIE [27], LR3M [36] and SDD [19]). It takes approximately 186.655s for the proposed method to process one image of size  $1312 \times 2000$ . MSR, PMSIRE, LIME, NPIE-MLLS and SDD require approximately 13.129s,

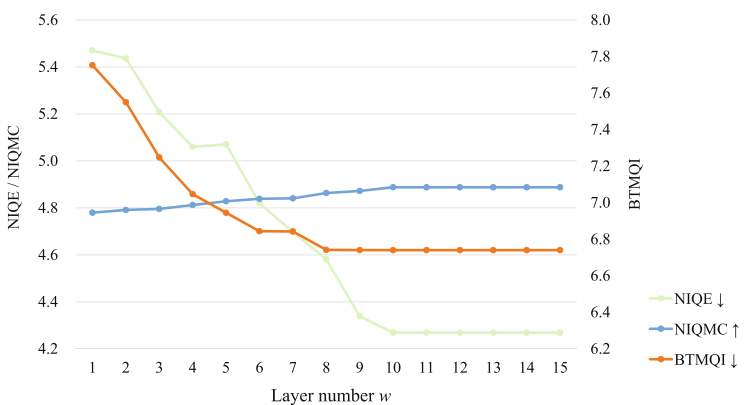
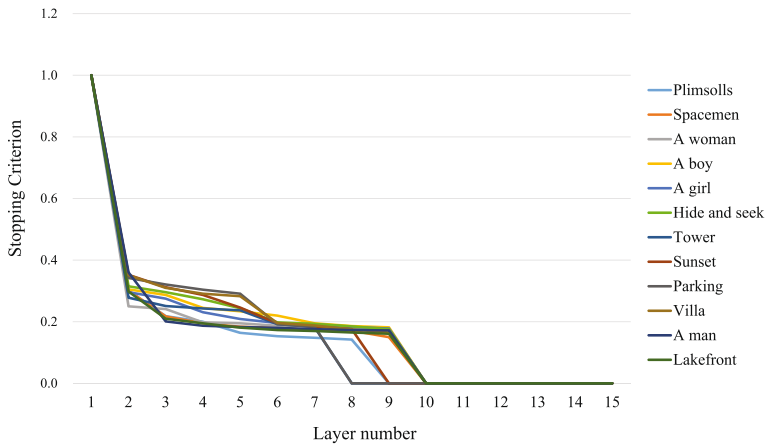


Fig. 15 Objective assessments with different layer numbers for *A man*



**Fig. 16** Convergence curves of the 12 example test images. The stopping criterion is when the difference of illumination or reflectance between adjacent layers is smaller than a threshold, *i.e.*,  $\|L_{i+1}(x, y) - L_i(x, y)\| / \|L_i(x, y)\| < 5 \times 10^{-3}$  (or  $\|R_{i+1}(x, y) - R_i(x, y)\| / \|R_i(x, y)\| < 5 \times 10^{-3}$ )

17.101s, 20.234s, 158.126s and 100.344s, respectively. Due to the computing complexity, SROLLIE and LR3M need a computer with larger memory to operate. To process an image of the same size, SROLLIE and LR3M require about 296.949 seconds and 2 hours, respectively. Even though our method needs more time than MSR, PMSIRE, LIME, NPIC-MLLS and SDD, the enhanced results of the proposed method are better than theirs in terms of subjective and objective assessments.

#### 4.5 Limitations and future work

Although the enhanced images of the proposed method are better than other methods, our technique still has a problem. That is, in some enhanced images, the details originally located in the extremely dark areas may be blurred, such as the red circle area in Fig. 5(l). This is because the contrast of extremely dark areas in the original image is not obvious, and HVS cannot perceive these regions well. This pinpoints a direction for our future works. To avoid ignorance of the scene content, semantic understanding is required. With further refinement, combining deep learning techniques with illumination decomposition will be performed. In addition, we will also focus on optimizing the proposed method in the future.

## 5 Conclusion

In this paper, a perceptive low-light image enhancement method via multi-layer illumination decomposition model is proposed. The contrast is defined from local and global aspects by taking noise into consideration. Then, the defined contrast is combined with the JND model to design the perceptive contrast, which can adapt to HVS and maintain the naturalness of the image for the subsequent enhancement. To estimate the illumination and reflectance components, a multi-layer illumination decomposition model is presented. This model can adaptively adjust the extent of contrast enhancement in the multi-scale adaptive filter based on the perceptive contrast. The proposed method can not only ensure the enhancement of brightness, but also preserve the naturalness of the image. The experimental results reveal

the advantages of our method compared with several state-of-the-arts in subjective assessment. Furthermore, BTMQI, NIQE and NIQMC objective assessments further demonstrate the superiority of the proposed method.

## References

1. Abdullah-Al-Wadud M, Kabir MH, Dewan MAA, CHae O (2007) A dynamic histogram equalization for image contrast enhancement. *IEEE Trans Consum Electron* 53(2):593–600
2. Cheng H, Long W, Li y, Liu H (2020) Two low illuminance image enhancement algorithms based on grey level mapping. *Multimed Tools Appl*
3. Chou CH, Li YC (1995) Perceptually tuned subband image coder based on the measure of just-noticeable-distortion profile. *IEEE Trans on Circuits & Systems for Video Technology* 5(6):467–476
4. Dabov K, Foi A, Katkovnik V, Egiazarian K (2007) Image denoising by sparse 3-d transform-domain collaborative filtering. *IEEE Trans Image Process* 16(8):2080–2095
5. Dicarolo JM, Wandell BA (2006) Rendering high dynamic range images. *Proc Spie* 3956:392–401
6. Dong X, Wang G, Pang Y, Li W, Wen J, Meng W, Lu Y (2011) Fast efficient algorithm for enhancement of low lighting video. In: 2011 IEEE international conference on multimedia and expo (ICME), pp 1–6
7. Edoardo P, Luca DC, Alessandro R, Daniele M (2005) Mathematical definition and analysis of the retinex algorithm. *J Opt Soc Am A: Opt Image Sci Vis* 22(12):2613–21
8. Eilertsen G, Mantiuk RK, Unger J (2015) Real-time noise-aware tone mapping. *ACM Trans Graph* 34(6):1–15
9. Fattal R, Lischinski D, Werman M (2002) Gradient domain high dynamic range compression. 21(3)
10. Feng X, Li J, Hua Z (2020) Low-light image enhancement algorithm based on an atmospheric physical model. *Multimed Tools Appl* 79(3)
11. Fu X, Liao Y, Zeng D, Huang Y, Zhang X, Ding X (2015) A probabilistic method for image enhancement with simultaneous illumination and reflectance estimation. *IEEE Trans Image Process* 24(12):4965–4977
12. Fu X, Zeng D, Huang Y, Liao Y, Ding X, Paisley J (2016) A fusion-based enhancing method for weakly illuminated images. *Signal Process* 129:82–96
13. Gonzalez RC, Woods RE (2007) *Digital Image Processing*, 3rd. Prentice-Hall, Upper Saddle River, NJ
14. Govind LP, Josemartin MJ (2019) Kerala Application of multi-stage filtering and multi-layer model in the context of dark and non uniformly illuminated images. In: 2019 2Nd international conference on intelligent computing, instrumentation and control technologies (ICICICT), vol 1, pp 615–620
15. Gu K, Wang S, Zhai G, Ma S, Yang X, Lin W, Zhang W, Gao W (2016) Blind quality assessment of tone-mapped images via analysis of information, naturalness, and structure. *IEEE Transactions on Multimedia* 18(3):432–443
16. Gu K, Lin W, Zhai G, Yang X, Zhang W, Chen CW (2017) No-reference quality metric of contrast-distorted images based on information maximization. *IEEE Trans Cybern* 47(12):4559–4565
17. Guo C, Li C, Guo J, Loy CC, Hou J, Kwong S, Cong R (2020) Zero-reference deep curve estimation for low-light image enhancement. In: 2020 IEEE/CVF conference on computer vision and pattern recognition (CVPR), pp 1777–1786
18. Guo X, Li Y, Ling H (2017) Lime: Low-light image enhancement via illumination map estimation. *IEEE Trans Image Process* 26(2):982–993
19. Hao S, Han X, Guo Y, Xu X, Wang M (2020) Low-light image enhancement with semi-decoupled decomposition. *IEEE Transactions on Multimedia* 22(12):3025–3038
20. Jayant N (1992) Signal compression: technology targets and research directions. *IEEE Journal on Selected Areas in Communications* 10(5):796–818
21. Jobson DJ, Rahman Z, Woodell GA (1997a) A multiscale retinex for bridging the gap between color images and the human observation of scenes. *IEEE Trans Image Process* 6(7):965–976
22. Jobson DJ, Rahman Z, Woodell GA (1997b) Properties and performance of a center/surround retinex. *IEEE Trans Image Process* 6(3):451–462
23. Kimmel R, Elad M, Shaked D (2003) Keshet r, A variational framework for retinex. *Int J Comput Vis, Sobel I*
24. Land EH (1977) The retinex theory of color vision. *Sci Am* 237(6):108–129
25. Lee C, Lee C, Kim C (2012) Contrast enhancement based on layered difference representation. In: 2012 IEEE international conference on image processing (ICIP), pp 965–968

26. Li L, Wang R, Wang W, Gao W (2015) A low-light image enhancement method for both denoising and contrast enlarging. In: 2015 IEEE international conference on image processing (ICIP), pp 3730–3734
27. Li M, Liu J, Yang W, Sun X, Guo Z (2018) Structure-revealing low-light image enhancement via robust retinex model. *IEEE Trans Image Process* 27(6):2828–2841
28. Liao S, Hu Y, Xiangyu Z, Li SZ (2015) Person re-identification by local maximal occurrence representation and metric learning. In: 2015 IEEE conference on computer vision and pattern recognition (CVPR), pp 2197–2206
29. Lim J, Heo M, Lee C, Kim CS (2017) Contrast enhancement of noisy low-light images based on structure-texture-noise decomposition. *J Vis Commun Image Represent* 45:107–121
30. Lore KG, Akintayo A, Sarkar S (2017) Llnet: a deep autoencoder approach to natural low-light image enhancement. *Pattern Recogn* 61:650–662
31. Ma K, Zeng K, Wang Z (2015) Perceptual quality assessment for multi-exposure image fusion. *IEEE Trans Image Process* 24(11):3345–3356
32. Mittal A, Soundararajan R, Bovik AC (2013) Making a “completely blind” image quality analyzer. *IEEE Signal Processing Letters* 20(3):209–212
33. Panetta KA, Wharton EJ, Aгаian SS (2008) Human visual system-based image enhancement and logarithmic contrast measure. *IEEE Trans Syst Man Cybern Part B (Cybern)* 38(1):174–188
34. Pisano ED, Zong S, Hemminger BM, Deluca M, Johnston RE, Muller K, Braeuning MP, Pizer SM (1998) Contrast limited adaptive histogram equalization image processing to improve the detection of simulated spiculations in dense mammograms. *J Digit Imaging* 11(4):193–200
35. Qiao Y, Liu Y, Yang X, Zhou D, Xu M, Zhang Q, Wei X (2020) Attention-guided hierarchical structure aggregation for image matting. In: 2020 IEEE/CVF conference on computer vision and pattern recognition (CVPR), pp 13673–13682
36. Ren X, Yang W, Cheng W, Liu J (2020) Lr3m: Robust low-light enhancement via low-rank regularized retinex model. *IEEE Trans Image Process* 29:5862–5876
37. Selim SZ, Ismail MA (1984) K-means-type algorithms: a generalized convergence theorem and characterization of local optimality. *IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-6(1):81–87*
38. Steyer S, Lenk C, Kellner D, Tanzmeister G, Wollherr D (2020) Grid-based object tracking with nonlinear dynamic state and shape estimation. *IEEE Trans Intell Transp Syst* 21(7):2874–2893
39. Wang C, Ye Z (2005) Brightness preserving histogram equalization with maximum entropy: a variational perspective. *IEEE Trans Consum Electron* 51(4):1326–1334
40. Wang D, Niu X, Dou Y (2014) A piecewise-based contrast enhancement framework for low lighting video. In: 2014 IEEE international conference on security, pattern analysis, and cybernetics (SPAC), pp 235–240
41. Wang LW, Liu ZS, Siu WC, Lun DPK (2020) Lightening network for low-light image enhancement. *IEEE Trans Image Process* 29:7984–7996
42. Wang R, Zhang Q, Fu C, Shen X, Zheng W, Jia J (2019) Underexposed photo enhancement using deep illumination estimation. In: 2019 IEEE/CVF conference on computer vision and pattern recognition (CVPR), pp 6842–6850
43. Wang S, Luo G (2018) Naturalness preserved image enhancement using a priori multi-layer lightness statistics. *IEEE Trans Image Process* 27(2):938–948
44. Wang S, Zheng J, Hu HM, Li B (2013) Naturalness preserved enhancement algorithm for non-uniform illumination images. *IEEE Trans Image Process* 22(9):3538–3548
45. Wu Y, Song W, Zheng J, Liu F (2021) Non-uniform low-light image enhancement via non-local similarity decomposition model. *Signal Processing Image Communication* 93(2):116141
46. Xu J, Hou Y, Ren D, Liu L, Zhu F, Yu M, Wang H, Shao L (2020) Star: a structure and texture aware retinex model. *IEEE Trans Image Process* 29:5022–5037
47. Xu K, Yang X, Yin B, Lau RWH (2020) Learning to restore low-light images via decomposition-and-enhancement. In: 2020 IEEE/CVF conference on computer vision and pattern recognition (CVPR), pp 2278–2287
48. Xu X, Luo X, Ma L (2020) Context-aware hierarchical feature attention network for multi-scale object detection. In: 2020 IEEE international conference on image processing (ICIP), pp 2011–2015
49. Xueyang F, Delu Z, Yue H, Xiaoping Z, Xinghao D (2016) A weighted variational model for simultaneous reflectance and illumination estimation. In: 2016 IEEE conference on computer vision and pattern recognition (CVPR), pp 2782–2790
50. Yang W, Wang S, Fang Y, Wang Y, Liu J (2020) From fidelity to perceptual quality: A semi-supervised approach for low-light image enhancement. In: 2020 IEEE/CVF conference on computer vision and pattern recognition (CVPR), pp 3060–3069



51. Yu L, Su H, Jung C (2018) Perceptually optimized enhancement of contrast and color in images. *IEEE Access* 6:36132–36142
52. Zhang C, Yan Q, Zhu Y, Li X, Sun J, Zhang Y (2020) Attention-based network for low-light image enhancement. [arXiv:2005.09829](https://arxiv.org/abs/2005.09829)
53. Zhang X, Shen P, Luo L, Zhang L, Song J (2012) Enhancement and noise reduction of very low light level images. In: 2012 IEEE international conference on pattern recognition (ICPR), pp 2034–2037
54. Zhang Y, Zhang J, Guo X (2019) Kindling the darkness: a practical low-light image enhancer. [arXiv:1905.04161](https://arxiv.org/abs/1905.04161)
55. Zhu M, Pan P, Chen W, Yang Y (2020) Eemefn: Low-light image enhancement via edge-enhanced multi-exposure fusion network. In: *AAAI Conference on artificial intelligence (AAAI)*, vol 34, pp 13106–13113

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.