



# Single-image low-light enhancement via generating and fusing multiple sources

Zhiyuan Zhou<sup>1</sup> · Zhuang Feng<sup>1</sup> · Jilong Liu<sup>1</sup> · Shijie Hao<sup>1</sup>

Received: 6 June 2018 / Accepted: 14 November 2018 / Published online: 24 November 2018  
© Springer-Verlag London Ltd., part of Springer Nature 2018

## Abstract

Imperfect lightness conditions usually lower the visual quality of an image by bringing in unclear image details and poor image contrast. Traditional low-light enhancement models based on one single input are often limited in avoiding the effect of over-enhancement or under-enhancement. Models based on fusing multiple input sources usually perform well in relieving this issue, as they can harmonize the complementary visual appearances of a same scene provided by different sources. Nevertheless, these models still have difficulty in dealing with the situation that only one input is at hand, which usually happens in many practical situations. In this paper, we propose a low-light enhancement model that artificially enriches input sources and then seamlessly fuses them. Specifically, with an input image, we first generate multiple enhanced images based on a lightness-aware camera response model. These images are then fused at mid-level based on a patch-based image decomposition model. To validate our model, we conduct qualitative and quantitative comparisons with several state-of-the-art single-source and multi-source models on a collection of real-world images. Experimental results show that our model better improves the image quality in terms of visual naturalness and aesthetics.

**Keywords** Low-light image · Image enhancement · Fusion

## 1 Introduction

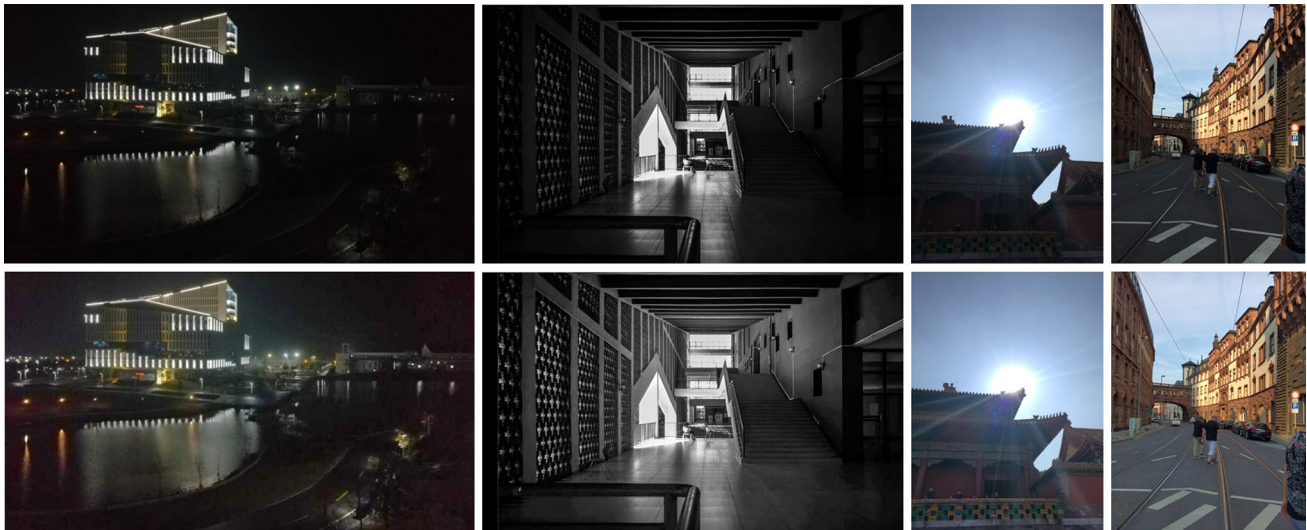
Nowadays, taking and sharing photographs has become a common life style in recent years. Various interesting image-based applications have emerged, such as image retrieval [1], classification [2, 3], clustering [4, 5], stylization [6], and segmentation [7]. Despite the advances in photographing devices, low-quality images can still be produced due to various reasons. On one hand, many people are not familiar with the basic shooting skills, e.g., rule of thirds, exposure setting, and so on. In this context, intelligent systems have been designed to help people choose a proper content composition [8, 9]. On the other hand, the visual appearance of a photograph can be affected by imperfect imaging conditions, such as low lightness, foggy weather, and motion blur. Therefore, enhancement models at the pixel level are also highly

desired. Typical applications include detail enhancement [10], color transfer [11], low-light enhancement [12], dehazing [13], motion deblur [14], to name but a few.

In this paper, we focus on solving the issue of low lightness, which is often encountered in the photograph shooting activities. We can generally divide the poor lightness into two types. The first one is the globally low lightness, where there is only a few weak light sources existing in the imaging scene, e.g., a nighttime outdoor place and a dimly lit indoor place in Fig. 1. The second one is the unbalanced lightness, where a good light source exists, but fails to well illuminate the whole scene, e.g., backlight and sidelight in Fig. 1. To address the first type of low lightness, many enhancing models have been proposed, such as histogram-based models [15–19], aiming at stretching the histogram to a larger range of intensity values. However, they are limited to the second type of unbalanced lightness conditions, as these models tend to over-enhance the originally well-illuminated regions. Retinex-based models [20–22] are able to relieve this issue to some extent, but they are still less effective in tackling complex lightness conditions. Commonly, all the above

✉ Shijie Hao  
shijie\_hao@hfut.edu.cn

<sup>1</sup> School of Computer and Information, Hefei University of Technology, Hefei City 230009, China



**Fig. 1** First row: Images with less lit regions under various conditions: nighttime, indoor, backlight, and sidelight. Second row: enhanced images based on our model

models try to improve the contrast via manipulating only one single input.

The advance in photographing device benefits the task of low-light enhancement by providing more input data. Specifically, a same scene can be recorded with multiple images with different exposures, which effectively expand the intensity range for an image, especially for the originally dark regions. The multiple sources pave the way for the development of fusion-based enhancing models [23–25]. However, these models depend on the enriched data sources, which can be still unavailable in many situations. For example, we only get access to only one single low-light image from the Internet in most cases. If we still use the road map of multi-source fusion, we have to artificially generate a few intermediate enhanced images as the inputs in advance.

Following this road map, only with one single low-light image, we propose a low-light enhancement model via generating and fusing multiple sources. With a lightness-aware camera response model, multiple initial enhancements simulating differently exposed images are firstly produced. They are then fused at mid-level based on a patch-based image representation, in which image patches from each source are decomposed into several signals and are fused, respectively. The final enhancement is obtained by recomposing the fused signals back again. The highlights of our research are twofold. First, we extend the ability of camera response function in terms of adapting to different lightness configurations. Second, the mid-level fusion show more competitive performance than the current state-of-the-art enhancing models, including the ones based on single source and multiple sources.

The rest of this paper is organized as follows. Section 2 introduces the related works. Section 3 presents the details

of the proposed model. Qualitative and quantitative comparisons are reported in Sect. 4. We finally conclude our research in Sect. 5.

## 2 Related works

In this section, we briefly review the related works of low-light image enhancement. We divided them into single-source models and multi-source models.

For the **single-source** group, a representative enhancing model is based on the manipulation of image histogram [15–19]. Based on the observation that the histogram of a low-light image heavily tails at low-intensity region, the histogram-based models mainly target at equalizing the intensity distribution across the whole intensity range [15–18] or reshaping the histogram into a desired distribution [19]. Since a histogram usually ignores most spatial information of an image, these models are often limited in tackling local lightness variation. As a result, they tend to produce the over-enhanced or under-enhanced effect.

Differently, Retinex-based models [20, 21, 26, 27] assume that an image is composed of an illumination layer and a reflectance layer. The former layer represents the illumination of the imaging scene, and the latter layer represents the inherent characteristics of object surface. A straightforward way for the Retinex-based models is to change the illumination layer, keep the reflectance layer unchanged, and recompose the two layers back again [26, 27]. The key component of Retinex-based enhancement models lies in the successful illumination–reflectance decomposition.

Since the image decomposition is ill-posed in nature, it often needs an alternative optimization process to

approximate the two layers, which can be unstable and time-consuming. Enhancement models based on simplified Retinex model are thereof proposed [22, 28]. These models still assume that an image is the combination of the two layers. The difference is that they roughly estimate the illumination layer with a simple MaxRGB technique and refine the MaxRGB image with an edge-preserving filter, which plays a vital role in these simplified Retinex models. For example, choosing a different filtering model can lead to slightly different enhancing effects, especially for image regions with complex patterns [28].

There is another interesting assumption that the inverse of a low-light image resembles a hazy image. By applying dehazing techniques, the darkness can be eliminated as haze in the reverse low-light image [29–31]. However, methods based on this assumption tend to generate unrealistic effect on salient object boundaries.

The single-source-based enhancement models are usually controlled by one or two parameters, e.g., (simplified) Retinex-based models, or completely parameter-free, e.g., histogram-based models, which act as a uniform enhancing strength imposed on the whole image region. Therefore, they are less spatially aware of image contents with different illumination conditions. In this context, they tend to produce improper local enhancing effects.

For the **multi-source** group, the above issue can be largely relieved by jointly considering multiple sources as the inputs for the enhancement model, which are potentially complementary to each other. With advanced imaging hardware, multiple source images of a same scene can be almost simultaneously collected with different exposures. The key left for the enhancement task is the seamless fusion of these input images, which adaptively combines the different appearances of a same image region, and thereof avoids the over-/under-enhancement. Bertalmio and Levine [24] propose to encode the gradient and color information from a short-exposed image and a long-exposed image into an image functional, and perform variational minimization to obtain the final fusion. Kou et al. [25] use a multi-resolution technique for achieving the seamless fusion. Additionally, they propose to improve the fused result by further enhancing image details. Specifically, they propose an improved image filter that extracts high-frequency details from multiple inputs and add them into the fused image. Ma et al. [23] propose a novel patch decomposition model that separates an image patch into three kinds of signals. Then, the decomposed signals from each source are linearly or nonlinearly fused. The resultant enhancement is finally obtained by recomposing the signals back again.

For the situation when only one image is at hand, the technical road map of multi-source fusion needs an expansion, and the stage of source generation becomes indispensable. In [32], along with the original input, Fu

et al. generate two intermediate enhanced images by applying two intensity transform techniques. Hao et al. [33] produce an intermediate enhancement based on the simplified Retinex model and fuse it with the original image by designing a content-aware weight map. Ying et al. [34] propose a novel bio-inspired enhancement model, in which the source is generated by a simulated camera response model [35]. Different from [34], the model in this paper avoids the heuristic judgment on if an image pixel is underexposed and thus is more flexible in generating more intermediate enhancing results.

Of note, there has been learning-based research for the low-light enhancement task [36, 37], which demonstrates very promising performance. For these methods, a collection of sufficient and reliable image pairs (normal lightness vs. low lightness) is vital to the training procedure.

## 3 Proposed method

### 3.1 Overall framework

Suppose the input  $\mathbf{I}_0 \in \mathbb{R}^{W \times H \times 3}$  is a color image represented in RGB space. The technical road map of our model is shown in Fig. 2. The model contains two main parts, i.e., lightness-aware source generation (described in Sect. 3.2) and multi-source fusion (described in Sect. 3.3).

Of note, the data flow in the first part is explained as follows. We first convert the RGB input into the HSV space and then only send the  $V$  channel into the generator. The reason is that the generator only aims at simulating different illumination conditions. After that, we replace the original  $V$  channel with the generated  $V$  channel and keep the other two channels unchanged. Then, for all the sources, we convert the HSV image back into the RGB space, which is used in the following stage of image decomposition.

### 3.2 Generation of fusion source

We generate the fusion sources based on the camera response model, which is jointly described by camera response function (CRF) and the brightness transform function (BTF). The former is only determined by a camera itself, while the latter is determined by the camera and the exposure ratio  $k$ . The general form of CRF can be represented as:

$$\mathbf{V} = f(\mathbf{E}). \quad (1)$$

Here,  $\mathbf{V} \in \mathbb{R}^{W \times H}$  is the  $V$  channel in the HSV space (observed pixel intensity of an image).  $\mathbf{E}$  is the ideal scene irradiance. By choosing exposure ratios, different observed images can be obtained, e.g.,  $\mathbf{V}_0 = f(\mathbf{E})$  (a trivial case as  $k = 1$ ),  $\mathbf{V}_1 = f(k\mathbf{E})$ .

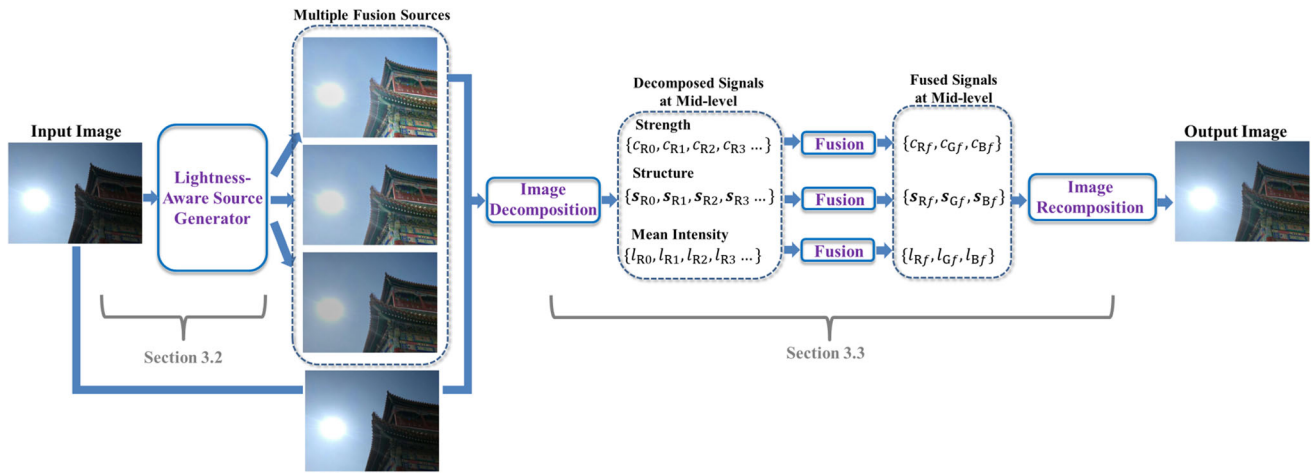


Fig. 2 Framework of our model

On the other hand, the mapping between  $\mathbf{V}_0$  and  $\mathbf{V}_1$  can be also described by the brightness transform function, which represents the mapping between two observed images of a same scene with different exposures:

$$\mathbf{V}_1 = g(\mathbf{V}_0, k) \tag{2}$$

Based on Eqs. 1 and 2, we have:

$$g(f(\mathbf{E}), k) = f(k\mathbf{E}) \tag{3}$$

According to [34, 35], we specify the BTF as a simple form:  $\mathbf{V}_1 = g(\mathbf{V}_0, k) = \beta \mathbf{V}_0^\gamma$ . Of note,  $\beta$  and  $\gamma$  are related to both the camera and the exposure ratio. Based on the comparametric equation [38], we have  $f(k\mathbf{E}) = \beta f(\mathbf{E})^\gamma$ . Except for the trivial case of  $\gamma = 1$ , we can obtain the closed form of CRF as:

$$f(\mathbf{E}) = e^{b(1-\mathbf{E}^a)}, \quad a = \log_k \gamma, \quad b = \frac{\ln \beta}{1-\gamma} \tag{4}$$

Here,  $a$  and  $b$  are build-in parameters of a camera. They can be empirically chosen as  $a = -0.3293$ ,  $b = 1.1258$ , which are suitable for most cameras [35]. Therefore, we can thus obtain the BTF parameters  $\beta$  and  $\gamma$ :

$$\beta = e^{b(1-k^a)}, \gamma = k^a \tag{5}$$

Then, the brightness transform function can be further parameterized by the exposure ratio  $k$ :

$$\mathbf{V}_1 = g(\mathbf{V}_0, k) = e^{b(1-k^a)} \mathbf{V}_0^{k^a} \tag{6}$$

In our application, we can take  $\mathbf{V}_0$  as the original image at hand and  $\mathbf{V}_1$  as a generated source.

Then, we estimate the exposure ratio  $k$  as follows. We first remove the small-scale image details from  $\mathbf{V}_0$  by using the fast guided filter [39]. Then, we extract the low-light regions of low lightness determined by a threshold  $\eta$ :

$$\mathbf{M}(\eta) = \{\mathbf{V}_0(p) | p < \eta\} \tag{7}$$

In another word,  $\mathbf{M}$  approximately indicates the set of low-light pixels in the original image. The exposure ratio estimation can be formed as an optimization problem:

$$\tilde{k}_\eta = \operatorname{argmax}_k \mathcal{H}(g(\mathbf{M}(\eta), k)), \tag{8}$$

where  $\mathcal{H}(\cdot)$  is the entropy and can be estimated from the image histogram of  $\mathbf{M}$ .

From the above modeling, the determination of low-light regions has large impact on the optimal exposure ratio. As exemplified in Fig. 1, low-light images can be divided into various specified conditions. Therefore, a single and ad hoc setting of the threshold  $\eta$  (e.g., 0.5 in [34]) is limited to describe the complex lightness conditions for an arbitrary image. We use a set of threshold values  $\{\eta_1, \eta_2, \dots, \eta_{N-1}\}$  to obtain different  $\tilde{k}$  values that cater to the generation of multiple sources for fusion.

### 3.3 Patch-based fusion

Without losing generality, we totally obtain  $N$  sources  $\{\mathbf{I}_0, \mathbf{I}_1, \dots, \mathbf{I}_{N-1}\}$  for the fusion process. For an image patch  $\mathbf{P}$  of each source, we adopt the patch-based image decomposition [23]:

$$\begin{aligned} \mathbf{P} &= \|\mathbf{P} - \mu_{\mathbf{P}}\| \cdot \frac{\mathbf{P} - \mu_{\mathbf{P}}}{\|\mathbf{P} - \mu_{\mathbf{P}}\|} + \mu_{\mathbf{P}} = \|\tilde{\mathbf{P}}\| \cdot \frac{\tilde{\mathbf{P}}}{\|\tilde{\mathbf{P}}\|} + \mu_{\mathbf{P}} \\ &= c \cdot \mathbf{s} + l \end{aligned} \tag{9}$$

In Eq. 9, for each  $M \times M$  squared patch, we stack its RGB channels together into a  $3M^2$ -length column vector  $\mathbf{P}$ . Here,  $\|\cdot\|$  is the L2 norm,  $c$  is the patch scale,  $\mathbf{s}$  is the patch structure, and  $l$  is the patch mean intensity. These decomposed elements can be seen as a mid-level representation of an image. In the following, the three components are separately fused.



**Table 1** Summary of our algorithm

Inputs: $\mathbf{I}_0, N, \{a, b\}, \{\eta_1, \eta_2, \dots, \eta_{N-1}\}, M, \rho, \sigma_g, \sigma_l$
# Source Generation
<b>Step 1:</b> Convert $\mathbf{I}_0$ into HSV space, and extract its V channel $\mathbf{V}_0$ . <i>for</i> $i=1:N-1$
<b>Step 2:</b> Estimate the optimal $\tilde{k}_{\eta_i}$ based on Eq.8.
<b>Step 3:</b> Generate the $i$ -th source $\mathbf{V}_i$ based on Eq.6.
<b>Step 4:</b> Convert $\{\mathbf{H}_0, \mathbf{S}_0, \mathbf{V}_i\}$ back into the RGB space as $\mathbf{I}_i$ . <i>end for</i>
# Patch-based Fusion
<i>for each patch position in</i> $\{\mathbf{I}_0, \mathbf{I}_1, \dots, \mathbf{I}_{N-1}\}$
<b>Step 1:</b> Stack its RGB channel pixels into a column vector $\mathbf{P}$ .
<b>Step 2:</b> Decompose the patches into $\{c_0, \mathbf{s}_0, l_0\}, \dots, \{c_{N-1}, \mathbf{s}_{N-1}, l_{N-1}\}$ based on Eq.9.
<b>Step 3:</b> Fuse the elements respectively based on Eq.10, Eq.11, and Eq.12.
<b>Step 4:</b> Recompose the fused patch based on Eq.14, and reform it into separated RGB channels. <i>end for</i>
<b>Step 5:</b> Obtain $\mathbf{I}_f$ by accumulating and normalizing all the overlapping reconstructed patches.
Output: $\mathbf{I}_f$

First, a nonlinear max-fusion is applied to the patch scale  $c_n$ :

$$\hat{c} = \max_{0 \leq n \leq N-1} c_n \tag{10}$$

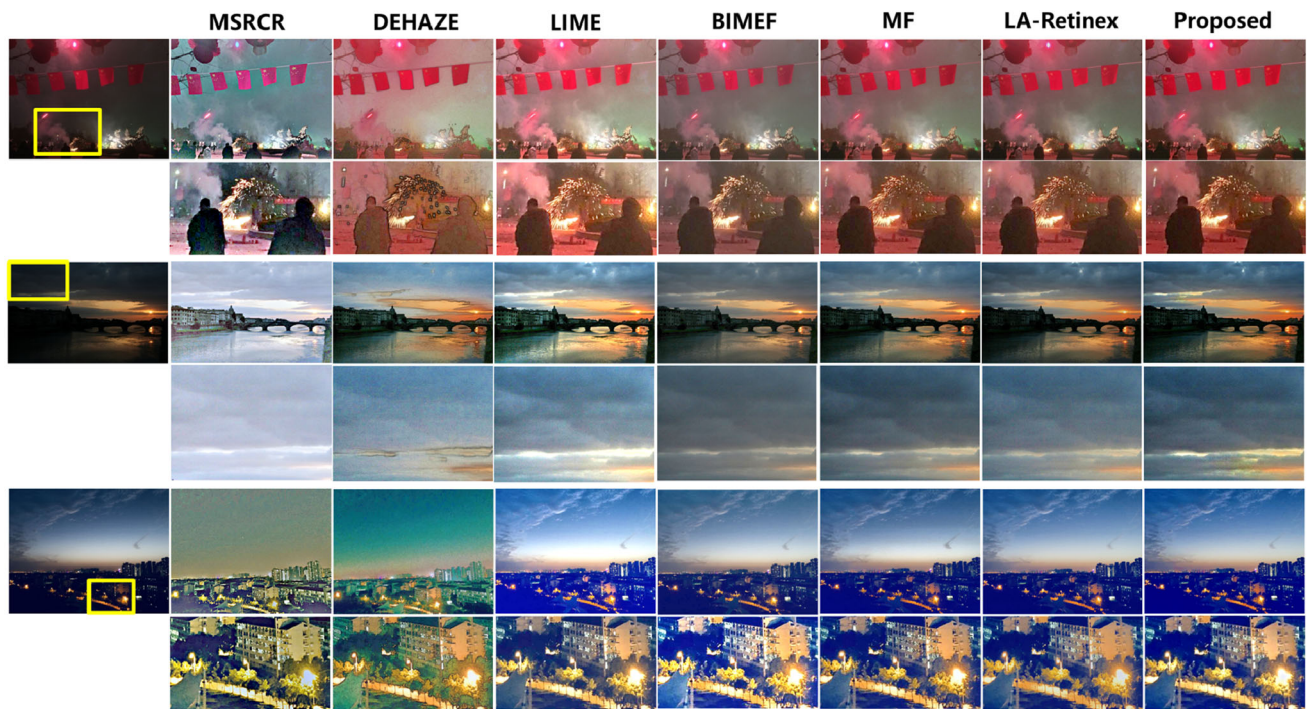
Second, a linear weight fusion is constructed for the patch structure  $\mathbf{s}_n$ :

$$\bar{\mathbf{s}} = \frac{\sum_{n=0}^{N-1} c_n^\rho \mathbf{s}_n}{\sum_{n=0}^{N-1} c_n^\rho} \tag{11}$$

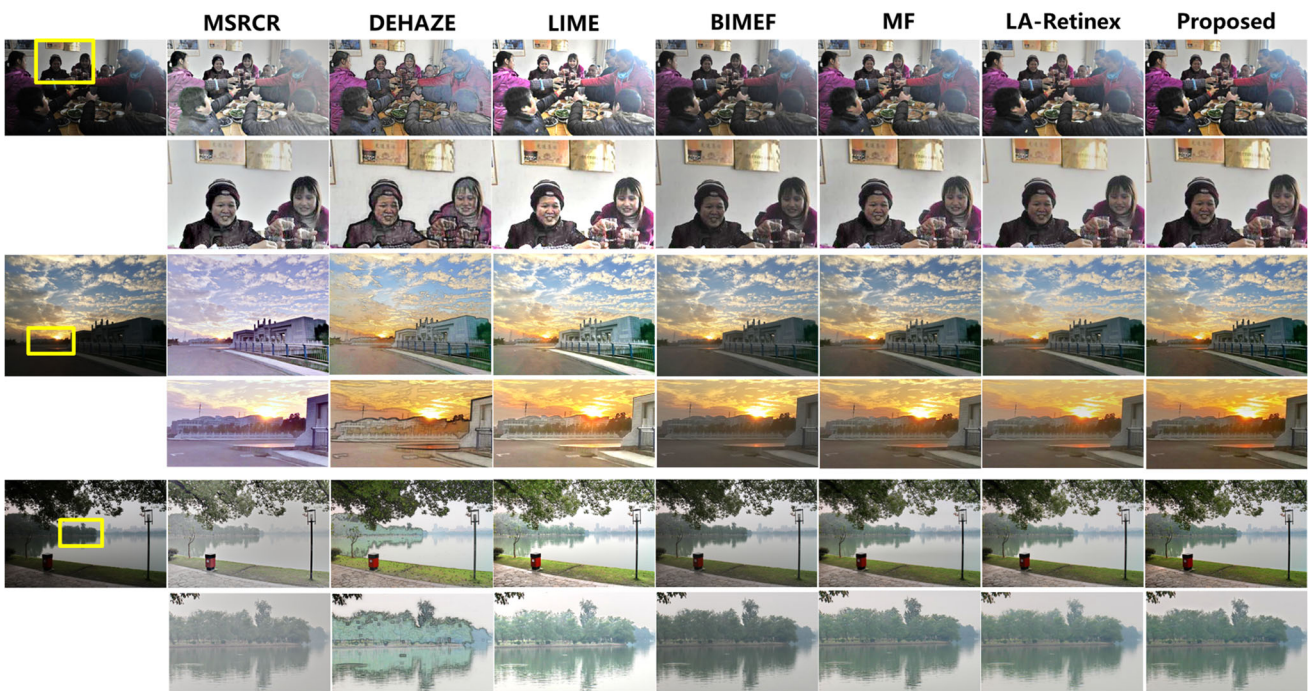
From Eq. 11, we observe that the fused  $\bar{\mathbf{s}}$  is jointly determined by  $\{\mathbf{s}_n\}$  of the multiple sources. The fusion weights are determined by the exponential of patch scales  $\{c_n\}$ , where  $\rho \geq 0$  is a hyper-parameter. The obtained  $\bar{\mathbf{s}}$  is further normalized as  $\hat{\mathbf{s}} = \bar{\mathbf{s}}/\|\bar{\mathbf{s}}\|$ . From Eqs. 10 and 11, we



**Fig. 3** A gallery of all experimental images



(a)



(b)

**Fig. 4** Visual comparison of the enhanced results based on all the seven models. (Better with an enlarged view and a bright screen display)

observe that the fusion tends to weigh more on strong patches and still considers the impact of weak patches. Third, we also use a weighted linear fusion for the patch mean  $l_n$ :

$$\hat{l} = \frac{\sum_{n=0}^{N-1} L(\mu_n^0, l_n) l_n}{\sum_{n=0}^{N-1} L(\mu_n^0, l_n)} \quad (12)$$

In Eq. 12,  $L(\cdot, \cdot)$  describes how well the lightness of  $l_n$  is in  $\mathbf{I}_n$ :



$$L(\mu_n^0, l_n) = \exp\left(-\frac{(\mu_n^0 - 0.5)^2}{2\sigma_g^2} - \frac{(l_n - 0.5)^2}{2\sigma_l^2}\right), \quad (13)$$

where  $\mu_n^0$  is the global mean of  $\mathbf{I}_n$ ,  $\sigma_g$  and  $\sigma_l$  control the spreads of the Gaussian distribution tails.

Finally, we disconnect the stacked  $\hat{\mathbf{s}}$  back into the RGB channels and reconstruct them with the obtained  $\hat{c}$  and  $\hat{l}$  according to Eq. 9:

$$\hat{\mathbf{P}}_\phi = \hat{c} \cdot \hat{\mathbf{s}}_\phi + \hat{l} \quad (14)$$

where  $\phi \in \{\mathbf{R}, \mathbf{G}, \mathbf{B}\}$  enumerates the three color channels. We use a sliding window with the stride length of  $B = \lfloor M/2 \rfloor$  to reconstruct each patch of the result image, and the pixels in overlapping regions are averaged. In this way, the reconstructed  $\mathbf{I}_f$  is taken as the final result. In our research, the fusion parameters  $M, \rho, \sigma_g, \sigma_l$  are typically set as in [23].

### 3.4 Algorithm summary

The whole algorithm is summarized in Table 1. We can easily see that the computational complexity of the whole algorithm is jointly determined by the total pixel number, number of fusion sources, and the patch size. In experiments, we empirically found that there exists a balance between the number of fusion sources and the computational efficiency. We choose  $N = 4$  sources for our research (including the original image) and set  $\eta$  as 0.4, 0.5, and 0.6 in our experiments.

## 4 Experiments

### 4.1 Experimental settings

In experiments, a total of 35 images were collected from the Internet or taken by the authors. As shown in Fig. 3, images taken at different outdoor and indoor scenes have various lightness conditions, e.g., nighttime, backlight, sidelight, and so on. We introduce six models for comparison: Single-source models include multi-scale Retinex-based model (MSRCR) [20], Dehazing-based model (DEHAZE) [30], LIME model (LIME) [22], while multi-source models include bio-inspired model (BIMEF) [34], multi-fusion model (MF) [32], lightness-aware simplified Retinex model (LA-Retinex) [33]. The codes of first five models were publicly available from the project webpage of [34], while the codes of [33] were implemented in our previous research. All of them were run on a laptop with 2.6G Hz CPU and 8G RAM.

**Table 2** Quantitative BTMQI scores [36]

	Input	MF	BIMEF	LA-Retinex	Proposed
1	6.5878	4.5735	4.9437	5.5423	<b>4.2190</b>
2	6.3720	4.0927	<b>3.9075</b>	4.4490	4.1746
3	7.2208	4.6509	5.5581	5.1587	<b>4.2323</b>
4	6.0125	<b>3.2141</b>	3.6453	3.9582	3.3125
5	6.1847	4.9536	6.0146	6.5771	<b>4.7097</b>
6	7.4190	3.1839	3.6623	3.6592	<b>3.0270</b>
7	5.4813	2.7359	3.8007	3.3349	<b>2.6670</b>
8	7.3200	<b>3.1428</b>	3.5406	3.7384	3.4612
9	5.3990	3.4674	4.4319	4.5723	<b>3.2071</b>
10	6.2633	5.3523	6.2086	5.9362	<b>5.2680</b>
11	6.2635	<b>5.3452</b>	5.4245	6.7638	5.9759
12	4.5570	2.4024	<b>1.9915</b>	2.2189	2.1526
13	4.5230	2.5870	2.4351	2.9699	<b>2.0999</b>
14	5.0670	2.6293	4.1233	3.3595	<b>2.5413</b>
15	4.9756	2.7189	3.0887	3.2874	<b>2.6776</b>
16	4.9923	3.7847	3.9701	4.1527	<b>3.3465</b>
17	2.7429	3.5937	3.5300	2.7628	<b>2.7230</b>
18	<b>2.3961</b>	4.3094	3.3738	3.7400	3.5146
19	<b>2.6597</b>	3.3691	3.2607	3.6066	2.7974
20	7.2735	<b>2.7968</b>	3.3818	4.4263	3.3552
21	5.7561	4.9557	5.4313	4.3126	<b>3.3926</b>
22	6.9872	4.5954	5.0727	5.5365	<b>4.5685</b>
23	6.9646	<b>4.8717</b>	5.4949	6.124	5.4167
24	6.6746	2.1805	<b>1.6504</b>	2.6584	2.0925
25	4.1777	<b>3.4732</b>	3.6408	4.8487	3.7679
26	5.9833	<b>2.4182</b>	2.8964	3.3035	2.5492
27	4.3502	2.3895	2.4301	2.8676	<b>2.3122</b>
28	3.2677	3.5747	<b>2.4506</b>	2.5768	2.7637
29	7.3812	<b>3.9731</b>	4.7241	5.7694	5.8768
30	3.6550	2.8941	2.5203	3.3041	<b>2.1831</b>
31	7.6865	7.1174	7.3682	7.6409	<b>6.9120</b>
32	5.0009	3.6898	4.5481	4.2073	<b>2.7662</b>
33	5.4931	2.7966	3.0804	3.8307	<b>2.4030</b>
34	5.7504	2.5032	2.8529	2.7184	<b>2.4364</b>
35	3.9162	4.1092	3.6193	<b>2.8941</b>	3.4462
1st num.	2	8	4	1	20
2nd num.	1	17	6	2	9

### 4.2 Visual comparisons

We first conduct visual comparisons. In Fig. 4a, we present the enhanced results of the three images with dim lightness. We have the following observations. First, all the models are able to reveal the image details hidden in the darkness, especially for the single-source models. Second, all the single-source models are prone to generating over-enhanced results, like the inappropriately boosted edges and

**Table 3** Quantitative aesthetic scores (color harmony, color vividness, and lightness) [37]

Input	MF	BIMEF	LA-Retinex	Proposed	
1	− 0.0360, − 0.4115, − 0.3933	0.4122, <i>0.2954</i> , − 0.0896	0.2842, − 0.0219, − 0.2109	<i>0.4293</i> , 0.1294, − <i>0.0803</i>	<b>0.4745, 0.3697</b> , − <b>0.0193</b>
2	0.4389, − 0.0004, 0.0661	<i>0.5277, 0.1326</i> , 0.2406	0.4996, 0.1246, 0.1814	0.5204, 0.1134, <i>0.2488</i>	<b>0.5641, 0.1606, 0.2651</b>
3	0.3454, 0.3671, − 0.1579	<i>0.5216</i> , 0.7047, <i>0.1538</i>	<b>0.5296</b> , 0.6230, 0.0977	0.4920, <i>0.7058</i> , 0.1478	0.5039, <b>0.7700, 0.1679</b>
4	0.3192, − 0.2669, − 0.1687	<b>0.6936</b> , <i>0.1215, 0.1381</i>	0.5499, 0.0227, 0.0155	0.5630, 0.0677, 0.0510	<i>0.6380</i> , <b>0.1252, 0.1524</b>
5	− 0.1217, − 0.7752, − 0.4543	<i>0.0262</i> , − <i>0.3318</i> , − <i>0.1414</i>	− 0.0609, − 0.5700, − 0.2805	− 0.0225, − 0.4072, − 0.1856	<b>0.0865</b> , − <b>0.2009</b> , − <b>0.0193</b>
6	0.0007, 0.4923, 0.2070	<b>0.2872</b> , <i>0.9973, 0.5317</i>	0.2368, 0.8536, 0.2984	0.2447, 0.9250, 0.4157	<i>0.2451, 1.0000, 0.4681</i>
7	− 0.2144, − 0.4635, − 0.3403	<b>0.0504</b> , − <i>0.2351</i> , − <i>0.0771</i>	− 0.1331, − 0.3301, − 0.3086	− 0.0416, − 0.3310, − 0.1831	<i>0.0314</i> , − <b>0.2308</b> , − <b>0.0723</b>
8	0.0382, − 0.3148, − 0.4133	<b>0.0978</b> , − <b>0.0258</b> , − <b>0.3926</b>	<i>0.0815</i> , − 0.1274, − 0.4810	0.0787, − 0.0821, − 0.4505	0.0532, − <i>0.0544</i> , − <i>0.4454</i>
9	− 0.1017, − 0.3245, − 0.3010	<i>0.2413, 0.0915</i> , − <b>0.0057</b>	0.1499, 0.0165, − 0.1977	<b>0.2649, 0.1219</b> , − 0.0440	0.2020, 0.0240, − <i>0.0422</i>
10	0.1509, 0.0347, 0.1447	<b>0.4665</b> , 0.4359, <i>0.4544</i>	0.4099, <b>0.4755</b> , 0.2446	0.3836, 0.4252, 0.4114	<i>0.4526, 0.4437, 0.5158</i>
11	0.1118, − 0.1360, − 0.0832	0.4023, <i>0.5281, 0.3824</i>	0.3444, 0.2804, 0.1781	<i>0.4156</i> , 0.4790, 0.3588	<b>0.4772, 0.5791, 0.4796</b>
12	− 0.1272, − 0.4043, − 0.3672	0.1220, − <i>0.1521</i> , − 0.3746	0.0136, − 0.2853, − 0.4149	<i>0.1482</i> , − 0.1785, − <i>0.3007</i>	<b>0.2440</b> , − <b>0.0477</b> , − <b>0.2605</b>
13	− 0.1952, − 0.9039, − 0.5887	<i>0.0718</i> , − <i>0.2114</i> , − <i>0.2037</i>	− 0.1529, − 0.4748, − 0.4784	− 0.0109, − 0.2308, − 0.2993	<b>0.1034</b> , − <b>0.1232</b> , − <b>0.1825</b>
14	0.3846, 0.6511, − 0.1326	<i>0.4330, 0.6516, 0.0045</i>	0.3724, 0.6000, − 0.0573	0.4112, 0.6195, − 0.0165	<b>0.4942, 0.7860, 0.0779</b>
15	− 0.0627, − 0.5050, − 0.3697	<b>0.4264</b> , − <i>0.0396</i> , − <i>0.0039</i>	0.2680, − 0.1340, − 0.0847	0.3550, − 0.0906, − 0.0438	<i>0.3893</i> , − <b>0.0202</b> , <b>0.0497</b>
16	0.2157, 0.4714, − 0.1165	<i>0.3510, 0.6728, 0.1280</i>	0.2701, 0.5352, − 0.0292	0.2823, 0.5688, − 0.0170	<b>0.3538, 0.7002, 0.1451</b>
17	0.1442, − 0.1937, − 0.3217	0.3828, − 0.1299, − 0.2192	0.3070, − <i>0.1221</i> , − 0.2381	<i>0.4632</i> , − 0.1299, − <i>0.1937</i>	<b>0.6234</b> , − <b>0.0917</b> , − <b>0.0402</b>
18	0.0286, − 0.3311, − 0.3357	<i>0.0847</i> , − <i>0.2191</i> , − <i>0.2953</i>	0.0647, − 0.3016, − 0.3693	0.0247, − 0.3672, − 0.3548	<b>0.2566</b> , − <b>0.1466</b> , − <b>0.2641</b>
19	0.5073, 0.1183, − 0.1033	<i>0.5798, 0.3806, 0.3160</i>	0.5453, 0.3102, 0.1889	0.5509, 0.2170, 0.2726	<b>0.6242, 0.3738, 0.3469</b>
20	− 0.0581, − 0.1223, 0.0906	<b>0.2503, 0.6079</b> , <i>0.3466</i>	0.1285, 0.3254, 0.1997	<i>0.2048, 0.5924, 0.4058</i>	0.1269, 0.3628, 0.2614
21	0.0368, − 0.3988, − 0.4124	0.4959, <i>0.1743</i> , − <i>0.2766</i>	<i>0.5246</i> , 0.1463, − 0.3554	0.5071, 0.1313, − 0.2824	<b>0.5457, 0.2300</b> , − <b>0.2296</b>
22	− 0.0241, − 0.6598, − 0.6724	0.0612, − 0.2784, − 0.6044	0.0406, − 0.3742, − 0.6377	<i>0.1127</i> , − <i>0.1560</i> , − <b>0.4879</b>	<b>0.1366</b> , − <b>0.1317</b> , − <i>0.4956</i>
23	− 0.0212, − 0.7583, − 0.5118	<i>0.1012</i> , − <i>0.5247</i> , − 0.3971	0.0238, − 0.6681, − 0.4562	0.0845, − 0.5520, − <i>0.3374</i>	<b>0.1323</b> , − <b>0.5090</b> , − <b>0.3152</b>
24	− <i>0.0116</i> , − <i>0.3208</i> , − 0.5269	− 0.0162, − 0.4020, − 0.1052	− 0.0502, − 0.4643, − 0.3605	− 0.0980, − 0.4642, − <i>0.0653</i>	<b>0.0690</b> , − <b>0.2603</b> , <b>0.1390</b>
25	0.1590, − 0.0207, − 0.2188	<i>0.3437, 0.4594</i> , − <i>0.0876</i>	0.3025, 0.2718, − 0.1123	0.3034, 0.2787, − 0.1040	<b>0.3353</b> , <i>0.4230</i> , − <b>0.0803</b>
26	0.0866, − 0.3156, − 0.4094	<i>0.4239, 0.1832</i> , − <i>0.1084</i>	0.3406, − 0.0220, − 0.2552	0.3849, 0.0790, − 0.1231	<b>0.4378</b> , <i>0.1330</i> , − <b>0.0076</b>
27	0.0493, − 0.7645, − 0.5315	<i>0.4012</i> , − <i>0.4433</i> , − <i>0.1075</i>	0.2632, − 0.5868, − 0.2892	0.3772, − 0.4815, − 0.1319	<b>0.4706</b> , − <b>0.3380</b> , − <b>0.0421</b>
28	0.2272, − 0.2115, − 0.4494	0.4312, − 0.1168, − 0.2382	0.3570, − 0.1805, − 0.3319	<i>0.4358</i> , − <i>0.0984</i> , − <b>0.2074</b>	<b>0.4520</b> , − <b>0.0854</b> , − 0.2389



**Table 3** (continued)

	Input	MF	BIMEF	LA-Retinex	Proposed
29	0.1423, 0.2290, – 0.1184	0.2622, <i>0.4236</i> , – 0.0426	<b>0.3123</b> , 0.3917, – 0.0598	0.2621, 0.3440, – 0.0206	<i>0.2715</i> , <b>0.4351</b> , <b>0.0272</b>
30	0.5654, 0.2904, – 0.0550	<i>0.6424</i> , 0.5505, 0.1107	0.6135, 0.4953, 0.0448	<b>0.6810</b> , <i>0.5677</i> , <i>0.1235</i>	0.6385, <b>0.5952</b> , <b>0.1535</b>
31	0.4830, – 0.1991, – 0.1809	<i>0.6342</i> , <i>0.0227</i> , <i>0.0924</i>	<b>0.6372</b> , – 0.0245, – 0.0218	0.4578, – 0.2018, – 0.0709	0.5669, <b>0.2685</b> , <b>0.2923</b>
32	0.1418, – 0.4914, – 0.4242	<i>0.5126</i> , <i>0.0356</i> , – 0.1458	0.3819, – 0.1339, – 0.2346	0.4480, – 0.0338, – 0.1599	<b>0.5639</b> , <b>0.0966</b> , – <b>0.0393</b>
33	0.2403, – 0.5509, – 0.4969	<i>0.6012</i> , <b>0.1863</b> , – 0.0296	0.5390, – 0.2188, – 0.2951	0.5971, <i>0.1675</i> , <i>0.0387</i>	<b>0.6063</b> , 0.1602, <b>0.0474</b>
34	0.0595, – 0.2644, 0.0329	<b>0.1856</b> , – <i>0.0753</i> , <i>0.1876</i>	<i>0.1692</i> , – <b>0.0576</b> , 0.1525	0.1012, – 0.1656, 0.1060	0.1399, – 0.0639, <b>0.2547</b>
35	– 0.1799, – 0.2760, – 0.4009	– 0.0103, – <i>0.0714</i> , – <i>0.3704</i>	– 0.0750, – 0.0918, – 0.3894	<i>0.0123</i> , – 0.1141, – 0.3958	<b>0.0241</b> , – <b>0.0295</b> , – <b>0.3540</b>
1st num.	0	17	5	5	78
2nd num.	2	61	4	23	15

textures (DEHAZE, LIME), or unrealistically change of global appearance (MSRCR). In contrary, the results based on multi-source models have much more balanced lightness configurations and are more visually appealing. In Fig. 4b, we present the enhanced results of the three images partially with low-light regions. We have similar observations as in Fig. 4a that multi-source models perform better than single-source ones. Furthermore, by taking a closer look, our model produces fewer artifacts and brings in more vivid colors. For example, our model has more natural appearance on the wall and medal region at the first row of Fig. 4b than MF and LA-Retinex. The color of grass, trees, and sunset region of our model is brighter than BIMEF in the second and third examples of Fig. 4b. The reasons are twofold. On one hand, the patch-based computation makes our method robust to artifacts to some extent. On the other hand, as the RGB channels of  $\{\mathbf{I}_0, \mathbf{I}_1, \dots, \mathbf{I}_{N-1}\}$  are jointly considered in the decomposition and fusion, our method is able to improve the color distribution.

### 4.3 Quantitative comparisons

We also conduct quantitative comparisons on all the models based on multi-source fusion. We use a non-reference image quality evaluator BTMQUI [40] and a visual aesthetic scoring network [41] and show all the scores in Tables 2 and 3. Of note, since our task does not change the image content composition, we only use the fine-grained scores of color harmony, color vividness, and lightness produced by the network trained in [41]. In the tables, we

use bold font/Italics to highlight the best/the second best performance, respectively.

From both tables, our model has the best performance among all four fusion-based models, which again validates the effectiveness of our model. Specifically, we also have some additional observations. First, the MF model has the second best performance. This confirms the usefulness of the fusion road map, in which good results can be obtained even by combining several simple enhancement models as in [32]. Second, for a few cases, the input images can have higher scores than those of their enhanced results. This observation indicates that enhancement does not necessarily improve the visual quality all the time. The reason is that the models for comparison are still not fully quality-aware or aesthetics-aware, although they try to harmonize the complementary appearances across the multiple sources with different fusion techniques.

## 5 Conclusions

In this paper, we propose a low-light enhancement model via generating and fusing multiple sources, which facilitates the situation that only one single input image is at hand. We empirically validate our model on various low-light images. Compared with single-source models and other multi-source models, our model produces better results in terms of visual naturalness and aesthetics. As mentioned above, although our model is able to improve the visual aesthetics of an image, it is still limited as the enhancement process itself is not aesthetics-aware. In the

following research, we plan to extend our model by equipping it with an aesthetics optimization process [42]. We also note that the determination of the low-light region in this paper is still heuristic to some extent. We can consider the technique of unsupervised feature selection [43–46] to accurately delineate the image regions with low lightness.

**Acknowledgements** The research was supported by the National Nature Science Foundation of China (Grant Nos. 61772171 and 61632007).

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

## References

- Wang M, Li H, Tao D, Lu K, Wu X (2012) Multimodal graph-based reranking for web image search. *IEEE Trans Image Process* 21(11):4649–4661
- Sun F, Tang J, Li H, Qi G, Huang T (2014) Multi-label image categorization with sparse factor representation. *IEEE Trans Image Process* 23(3):1028–1037
- Zhang S, Li X, Zong M, Zhu X, Wang R (2018) Efficient kNN classification with different numbers of nearest neighbors. *IEEE Trans Neural Netw Learn Syst* 29(5):1774–1785
- Zhu X, Zhang S, Li Y, Zhang J, Yang L, Fang Y (2018) Low-rank Sparse Subspace for spectral clustering. *IEEE Trans Knowl Data Eng.* <https://doi.org/10.1109/TKDE.2018.2858782>
- Hong R, Zhang L, Tao D (2016) Unified photo enhancement by discovering aesthetic communities from flickr. *IEEE Trans Image Process* 25(3):1124–1135
- Wang M, Hong R, Yuan X, Yan S, Chua T (2012) Movie2-Comics: towards a lively video content presentation. *IEEE Trans Multimed* 14(3–2):858–870
- Guo Y, Chen T (2018) Semantic segmentation of RGBD images based on deep depth regression. *Pattern Recogn Lett* 109:55–64
- Yin W, Mei T, Chen C, Li S (2014) Socialized mobile photography: learning to photograph with social context via mobile devices. *IEEE Trans Multimed* 16(1):184–200
- Ni B, Xu M, Wang M, Yan S, Tian Q (2013) Learning to photograph: a compositional perspective. *IEEE Trans Multimed* 15(5):1138–1151
- Hao S, Pan D, Guo Y, Hong R, Wang M (2016) Image detail enhancement with spatially guided filters. *Sig Process* 120:789–796
- Nguyen R, Kim S, Brown M (2014) Illuminant aware gamut-based color transfer. *Comput Graph Forum* 33(7):319–328
- Guo X, Li Y, Ling H (2017) LIME: low-light image enhancement via illumination map estimation. *IEEE Trans Image Process* 26(2):982–993
- Li Y, Tan R, Brown M (2015) Nighttime haze removal with glow and multiple light colors. In: *Proceedings of international conference on computer vision (ICCV)*
- Pan J, Hu Z, Su Z, Yang M (2017) L0-regularized intensity and gradient prior for deblurring text images and beyond. *IEEE Trans Pattern Anal Mach Intell* 39(2):342–355
- Reza AM (2004) Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement. *Journal VLSI Signal Process Syst Signal Image Video Technol* 38(1):35–44
- Lee C, Lee C, Kim C (2013) Contrast enhancement based on layered difference representation of 2D histograms. *IEEE Trans Image Process* 22(12):5372–5384
- Wang C, Ye Z (2005) Brightness preserving histogram equalization with maximum entropy: a variational perspective. *IEEE Trans Consum Electron* 51(4):1326–1334
- Chen S, Ramli A (2003) Minimum mean brightness error bi-histogram equalization in contrast enhancement. *IEEE Trans Image Process* 49(4):162–174
- Li Y, Guo F, Tan R, Brown M (2014) A contrast enhancement framework with JPEG artifacts suppression. In: *Proceedings of European Conference on Computer Vision (ECCV)*
- Jobson J, Rahman U, Woodell A (1997) A multi-scale retinex for bridging the gap between color images and the human observation of scenes. *IEEE Trans Image Process* 6(7):965–976
- Yue H, Yang J, Sun X, Wu F, Hou C (2017) Contrast enhancement based on intrinsic image decomposition. *IEEE Trans Image Process* 26(8):3981–3994
- Guo X, Li Y, Ling H (2017) LIME: low-light image enhancement via illumination map estimation. *IEEE Trans Image Process* 26(2):982–993
- Ma K, Li H, Yong H, Wang Z, Meng D, Zhang L (2017) Robust multi-exposure image fusion: a structural patch decomposition approach. *IEEE Trans Image Process* 26(5):2519–2532
- Bertalmio B, Levine S (2013) Variational approach for the fusion of exposure bracketed pairs. *IEEE Trans Image Process* 22(2):712–723
- Kou F, Wei Z, Chen W, Wu X, Wen C, Li Z (2018) Intelligent detail enhancement for exposure fusion. *IEEE Trans Multimedia* 20(2):484–495
- Fu X, Zeng D, Huang Y, Zhang X, Ding X (2016) A weighted variational model for simultaneous reflectance and illumination estimation. In: *Proceedings of computer vision and pattern recognition (CVPR)*
- Cai B, Xu X, Guo K, Jia K, Hu B, Tao D (2017) A joint intrinsic-extrinsic prior model for retinex. In: *Proceedings of international conference on computer vision (ICCV)*
- Hao S, Feng Z, Guo Y (2018) Low-light image enhancement with a refined illumination map. *Multim Tools Appl.* <https://doi.org/10.1007/s11042-017-5448-5>
- Malm H, Oskarsson M, Warrant E, Clarberg P, Hasselgren J, Lejdfors C (2007) Adaptive enhancement and noise reduction in very low light-level video. In: *Proceedings of international conference on computer vision (ICCV)*
- Dong X, Wang G, Pang Y (2011) Fast efficient algorithm for enhancement of low lighting video. In: *Proceedings of international conference on multimedia & expo (ICME)*
- Song J, Zhang L, Shen P, Peng X, Zhu G (2016) Single low-light image enhancement using luminance map. In: *Proceedings of Chinese conference of pattern recognition (CCPR)*
- 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
- Hao S, Guo Y, Wei Z (2018) Lightness-aware contrast enhancement for images with different illumination conditions. *Multimed Tools Appl.* <https://doi.org/10.1007/s11042-018-6257-1>
- Ying Z, Li G, Gao W (2017) A bio-inspired multi-exposure fusion framework for low-light image enhancement. [ArXiv:abs/1711.00591](https://arxiv.org/abs/1711.00591)
- Ying Z, Li G, Ren Y, Wang R, Wang W (2017) A new low-light image enhancement algorithm using camera response model. In: *Proceedings of international conference on computer vision (ICCV) workshop*

36. Lore K, Akintayo A, Sarkar S (2017) LLNet: a deep autoencoder approach to natural low-light image enhancement. *Pattern Recognit* 61:650–662
37. Chen C, Chen Q, Xu J, Koltun V (2018) Learning to see in the dark. In: *Proceedings of computer vision and pattern recognition (CVPR)*
38. Steve M (2000) Comparametric equations with practical applications in quantigraphic image processing. *IEEE Trans Image Process* 9(8):1389–1406
39. He K, Sun J (2015) Fast guided filter. *ArXiv*, abs/1505.00996
40. 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 Trans Multimed* 18(3):432–443
41. Kong S, Shen X, Lin Z, Mech R, Fowlkes C (2016) Photo aesthetics ranking network with attributes and content adaptation. In: *Proceedings of European conference on computer vision (ECCV)*
42. Zhang L, Wang M, Nie L, Hong R, Xia Y, Zimmermann R (2015) Biologically inspired media quality modeling. In: *Proceedings of ACM multimedia (ACM MM)*
43. Zheng W, Zhu X, Wen G, Zhu Y, Yu H, Gan J (2018) Unsupervised feature selection by self-paced learning regularization. *Pattern Recogn Lett*. <https://doi.org/10.1016/j.patrec.2018.06.029>
44. Zhu X, Zhang S, Hu R, Zhu Y, Song J (2018) Local and global structure preservation for robust unsupervised spectral feature selection. *IEEE Trans Knowl Data Eng* 30(3):517–529
45. Zheng W, Zhu X, Zhu Y, Hu R, Lei C (2018) Dynamic graph learning for spectral feature selection. *Multimed Tools Appl*. <https://doi.org/10.1007/s11042-017-5272-y>
46. Lei C, Zhu X (2018) Unsupervised feature selection via local structure learning and sparse learning. *Multimed Tools Appl*. <https://doi.org/10.1007/s11042-017-5381-7>

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