



Enhancing Low-Light Images with JPEG Artifact Based on Image Decomposition

Chenmin Xu, Shijie Hao^(✉), Yanrong Guo, and Richang Hong

Hefei University of Technology, Hefei, China
hfut.hsj@gmail.com

Abstract. Images shared on the Internet are often compressed into a small size, and thus have the JPEG artifact. This issue becomes challenging for task of low-light image enhancement, as the artifacts hidden in dark image regions can be further boosted by traditional enhancement models. We use a divide-and-conquer strategy to tackle this problem. Specifically, we decompose an input image into an illumination layer and a reflectance layer, which decouple the issues of low lightness and JPEG artifacts. Therefore we can deal with them separately with the off-the-shelf enhancing and deblocking techniques. Qualitative and quantitative comparisons validate the effectiveness of our method.

Keywords: Image enhancement · Low light · JPEG artifact · Retinex model

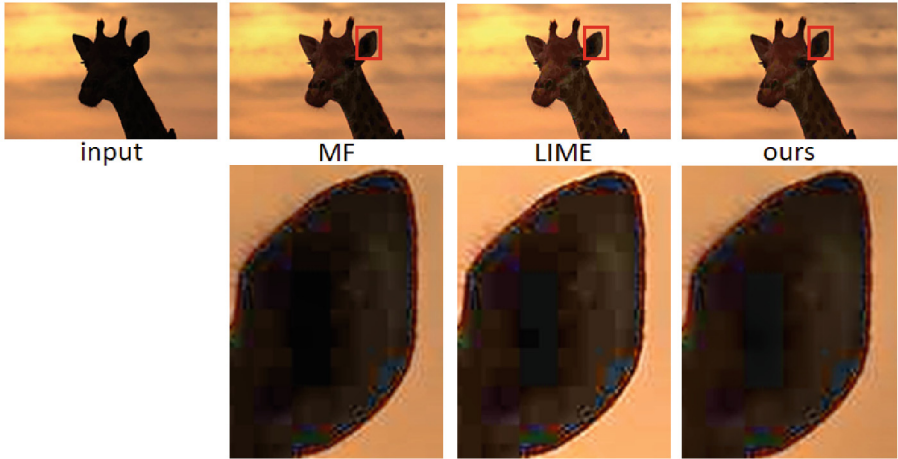
1 Introduction

With the advance of photographing devices, people always prefer images with clear content details and few artifacts. Nevertheless, this is not always the case in various real-world situations. For example, people nowadays enjoy taking photographs and share them on the Internet. In this process, the quality of an image can be affected by many factors. In the stage of photograph shooting, poor lightness conditions (e.g. nighttime) and amateur shooting skills (e.g. back light) often bring in a dark visual appearance ((e.g. Fig. 1(a))). In the stage of photograph distribution, images are often unintentionally compressed or resized into a smaller size by social network software like WeChat or QQ. Many high-frequency image details can be filtered out in this process and JPEG artifacts (or called block artifacts) are therefore introduced. In this context, for these compressed low-light images, an image enhancement model is expected to have the abilities of lightness enhancement and block removal [1].

However, conventional image enhancement methods [2–5] are limited due to the following reasons. First, these methods are not equipped with the artifact-removing ability. When we directly apply the off-the-shelf low-light enhancement methods on them, the JPEG artifacts can be unnecessarily amplified (e.g. Fig. 1(b)). Furthermore, JPEG artifacts usually hide in the image regions of low contrast, which makes the enhancement task more challenging. For example, compared with the original clean image, the hidden JPEG artifacts are not visually obvious in the compressed images (e.g. Fig. 1(a)).



(a) A visual comparison between low-light images without (first row) and with (second row) JPEG artifacts (Q=60). We use these six compressed low-light images as our experimental data.



(b) Enhanced results directly based on state-of-the-art low-light enhancement methods (e.g. MF [12] and LIME [5]), and our method (better with a bright screen display).



(c) Enhanced results based on the technical roadmaps of deblock-enhance, enhance-deblock, and our method (better with a bright screen display).

Fig. 1. An illustration of our research background.

Intuitively, we can simply adopt the technical roadmap of deblock-enhance or alternatively enhance-deblock to tackle this problem. However, both of them are still limited. As shown in Fig. 1(c), the deblock-enhance roadmap tends to produce over-smoothed final results, while the enhance-deblock roadmap has difficulties in removing the unnecessarily boosted JPEG artifact. In this paper, we propose a novel framework that simultaneously enhances the lightness and removes the JPEG blocks well (Fig. 2). The key element is to decompose the input image into the illumination layer and the reflectance layer. In this way, the low lightness and JPEG artifact can be well separated, which avoids the risk of artifact-boosting. Our research is highlighted in the proposed framework that well enhances compressed low-light images. Experimental results demonstrate the superior performance over other related methods both qualitatively and quantitatively.

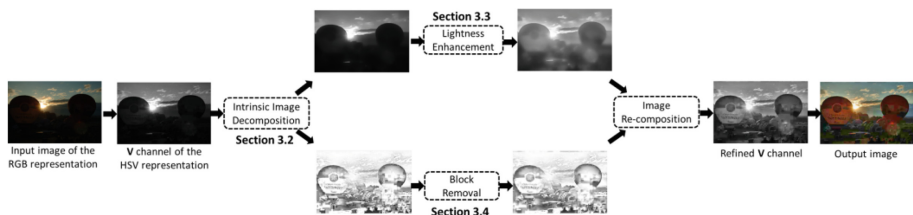


Fig. 2. The flowchart of our method.

2 Related Works

In recent years, many low-light image enhancement methods have been proposed. We can divide the enhancing methods into the single-source group and the multi-source group. The former refers to the methods with a single input image for enhancement, while the later uses multiple images as inputs.

As for single-source low-light image enhancement methods, we further divided into the histogram-based ones and the Retinex-based ones. As the histogram of a low-light image is heavily tailed at low intensities, the histogram-based methods are targeted on reshaping the histogram distribution [2, 6]. Since an image histogram is a global descriptor and drops all the spatial information, these methods are prone to produce over-enhanced [2] or under-enhanced [6] results for local regions. The Retinex-based models assume that an image is composed of an illumination layer representing the light intensity of an object, and a reflectance layer representing the physical characteristics of the object’s surface. Based on this image representation, the low-light enhancement can be achieved by adjusting the illumination layer [3, 7, 8]. Differently, Guo et al. [5] propose a simplified Retinex model for the task of low-light image enhancement. Instead of the intrinsic decomposition, they directly estimate a piece-smooth map as the illumination layer. In general, the visual appearance of these single-source methods’ results heavily depends on a properly chosen enhancing strength.

Low-light image enhancement methods based on multiple sources can relieve the issue of choosing the enhancing parameter, as they adopt the technical roadmap of multi-source fusion. In principle, larger dynamic range can be captured for an imaging scene with multiple images of different exposures. Then these images are fused based on a multi-scale image pyramid [9] or a patch-based image composition [10]. In many cases, however, multiple sources are not available and only one low-quality input image is at hand. To address this challenge, a feasible way is to artificially produce multiple initial enhancements and then fuse them. Ying et al. [12] propose a method simulating the camera response model and generate multiple intermediate images of different exposing levels. Different from [11], multiple initial results with different enhancing models are generated and taken as the fusion sources in [12].

Although all the above methods well address the problem of dark image appearance, they are not equipped with the function of artifact removal. To the best of our knowledge, there are few works that concentrate on the enhancement of compressed low-light images. Li et al. [1] propose to decompose an image into the structure layer and a texture layer. The former layer is for the contrast enhancement; while the later one is for the JPEG block removal. Our method resembles the backbone of [1] but distinguishes itself in the technical details. First, we use a totally different image representation model (illumination-reflectance vs. structure-texture). Second, our enhancement model is also different from the one adopted in [1] (fusion-based vs. histogram-based).

3 Proposed Method

3.1 Overall Framework

Our technical roadmap is to separately solve the low lightness issue and the JPEG artifact issue. The proposed framework is shown in Fig. 2. We first convert the RGB input image into the HSV space. Then the V channel \mathbf{V} of the input image \mathbf{S} is firstly decomposed into two components (Sect. 3.2), i.e. the illumination layer \mathbf{I} and the reflectance layer \mathbf{R} . Then we perform low-light enhancement on the illumination layer (Sect. 3.3), and perform the JPEG artifact removal on the reflectance layer (Sect. 3.4). Finally, the output image is obtained by re-combining the refined \mathbf{I}' and \mathbf{R}' : $\mathbf{V}_{\text{output}} = \mathbf{I}' \cdot \mathbf{R}'$. By replacing the refined $\mathbf{V}_{\text{output}}$ with the original \mathbf{V} , the final output image $\mathbf{S}_{\text{output}}$ can be obtained by converting the HSV representation back into the RGB representation. In another word, we keep the color information of \mathbf{S} unchanged during the whole process.

3.2 Image Decomposition

We decompose the input image \mathbf{V} into the illumination layer \mathbf{I} and the reflectance layer \mathbf{R} . Since we only have an observed image \mathbf{S} at hand in real-world applications, the decomposition is an ill-posed task. Additionally, for our task, we further aim to decouple the low lightness and the JPEG artifact.

To meet the above demands, we use an image decomposition model by jointly considering shape, texture and illumination priors [8]. We minimize the following target function:

$$E(\mathbf{I}, \mathbf{R}) = \|\mathbf{V} - \mathbf{I} \cdot \mathbf{R}\|_2^2 + \eta_1 E_s(\mathbf{I}) + \eta_2 E_t(\mathbf{R}) + \eta_3 E_l(\mathbf{I}) \quad (1)$$

Here the first term represents the data fidelity, and the rest terms encode the three priors. The first prior is constructed as:

$$E_s(\mathbf{I}) = u_x \|\nabla_x \mathbf{I}\|_2^2 + u_y \|\nabla_y \mathbf{I}\|_2^2 \quad (2)$$

$$u_x = \left(\frac{1}{\Omega} \sum_{\Omega} \nabla_x \mathbf{I} \left| |\nabla_x \mathbf{I}| + \varepsilon \right. \right)^{-1}, \quad u_y = \left(\frac{1}{\Omega} \sum_{\Omega} \nabla_y \mathbf{I} \left| |\nabla_y \mathbf{I}| + \varepsilon \right. \right)^{-1} \quad (3)$$

The minimization of E_s leads to the extracted \mathbf{I} that is consistent with the image structures of \mathbf{V} . The second prior is constructed as:

$$E_t(\mathbf{R}) = v_x \|\nabla_x \mathbf{R}\|_2^2 + v_y \|\nabla_y \mathbf{R}\|_2^2 \quad (4)$$

$$v_x = (|\nabla_x \mathbf{R}| + \varepsilon)^{-1}, \quad v_y = (|\nabla_y \mathbf{R}| + \varepsilon)^{-1} \quad (5)$$

The minimization of E_t preserves fine details of \mathbf{V} for the extracted \mathbf{R} . The third prior is constructed as

$$E_l(\mathbf{I}) = \|\mathbf{I} - \mathbf{B}\|_2^2 \quad (6)$$

where \mathbf{B} is the maxRGB matrix:

$$\mathbf{B}(p) = \max_{c \in \{\mathbf{R}, \mathbf{G}, \mathbf{B}\}} S_c(p) \quad (7)$$

The element in \mathbf{B} represents the possible maximum brightness of a pixel. Through the optimization, the obtained \mathbf{I} is forced to be consistent with the brightness distribution. An alternative optimization strategy is then applied to solve the optimal \mathbf{I} and \mathbf{R} .

In summary, the first prior E_s and the third prior E_l enforce the illumination layer to be structure-aware and illumination-aware; the second prior E_t focuses on preserving as many texture details and block artifacts as possible for the reflectance layer. The examples in Fig. 3 validate the effectiveness of the chosen decomposition model. Specifically, we observe that the decomposed \mathbf{R} contains almost all the block artifacts.

3.3 Low-Light Enhancement

We enhance the illumination layer \mathbf{I} by using a fusion framework proposed in [8]. As we only have \mathbf{I} at hand, we have to artificially produce several fusion sources. By noticing that the extracted \mathbf{I} has already been piece-wisely smooth and detail-free,

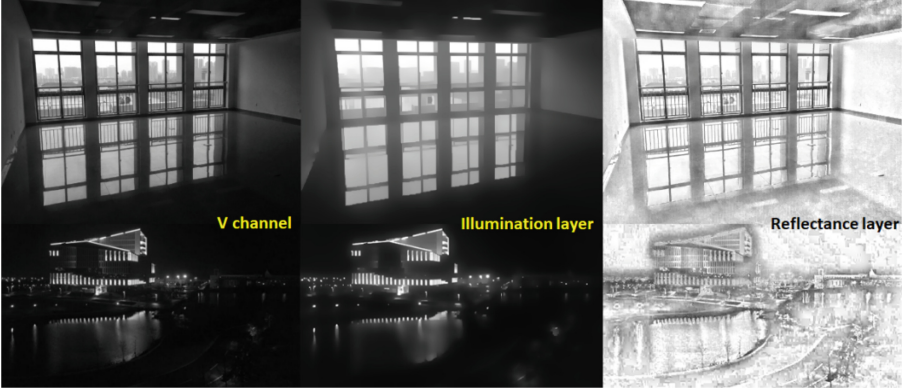


Fig. 3. Examples of the image decomposition results.

we can simply apply the global contrast enhancement models. For producing the first fusion source \mathbf{I}_1 , we first use a nonlinear intensity mapping:

$$\mathbf{I}_1(p) = \frac{2}{\pi} \arctan(\lambda \mathbf{I}(p)) \quad (8)$$

$$\lambda = \frac{1 - \text{mean}(\mathbf{I})}{\text{mean}(\mathbf{I})} + 10 \quad (9)$$

Then the enhancement result based on the well-known CLAHE method [3] is taken as the second fusion source \mathbf{I}_2 . In case of over- or under-enhancement, we choose the original \mathbf{I} as the third fusion source \mathbf{I}_3 , which plays the role of regularization in the fusion process.

We construct pixel-level weight matrices. We first consider the brightness of $\mathbf{I}_k (k = 1, 2, 3)$:

$$\mathbf{W}_B^k(p) = \exp - \frac{(\mathbf{I}_k(p) - \beta)^2}{2\sigma^2} \quad (10)$$

where β and σ represent the mean and standard deviation of the brightness in a natural image in a broadly statistical sense. They are empirically set as 0.5 and 0.25, respectively. When the pixel intensities are far from β , it means that they are possibly over- or under exposed, and the weights should be small. Second, we consider the chromatic contrast by incorporating the H and S channels of \mathbf{S} :

$$\mathbf{W}_C^k(p) = \mathbf{I}_k(p) \cdot (1 + \cos(\alpha \mathbf{H}(p) + \varphi) \cdot \mathbf{S}(p)) \quad (11)$$

where α and φ are parameters to preserve the color consistency, and empirically set as 2 and 1.39π . This weight emphasizes the image regions of high contrast and good colors.

By combining these two weights together and normalizing it, we can obtain the weights for each fusion source:

$$\bar{W}_k(p) = \frac{W_k(p)}{\sum_k W_k(p)} \quad (12)$$

$$W_k(p) = W_B^k(p) \cdot W_C^k(p) \quad (13)$$

To ensure a seamless fusion, we use multi-scale technique based on image pyramids. We first build Laplacian pyramids $L_l\{\mathbf{I}_k\}$ for all the fusion sources. By building Gaussian pyramids $G_l\{\bar{W}_k\}$ for the weight matrices, we can fuse $\{\mathbf{I}_k\}$ at various scales:

$$L'_l = \sum_k G_l\{\bar{W}_k\} L_l\{\mathbf{I}_k\} \quad (14)$$

At last, the enhanced illumination layer \mathbf{I}' can be obtained by collapsing the fused pyramid.

3.4 JPEG Artifact Removal

Since the JPEG artifact is divided from the illumination layer \mathbf{I} , we can conduct the deblocking in the reflectance layer \mathbf{R} . We adopt a simple but effective deblocking model in [1]:

$$\min_{R'(p)} \sum_p (R'(p) - R(p))^2 + \mu \sum_{p' \in \mathcal{N}(p)} (\nabla R'(p'))^2 \quad (15)$$

In this model, the first fidelity term restricts the refined \mathbf{R}' from going too far from the original \mathbf{R} . The second term is specifically designed for eliminating block edges that are introduced by JPEG compression. By considering the specific pattern of JPEG blocks, we choose a specific neighboring system $\mathcal{N}(p)$ for each location p : $p' \in \mathcal{N}(p)$ refers to pixel positions along the border edges of each 8×8 image patch. By this setting, the optimization process of Eq. 15 concentrates on the JPEG blocks, and tries to preserve the original image structures. The parameter μ acts as a balancing weight between the two terms, and is empirically set as 0.5.

4 Experiments

In this section, we validate our method with qualitative and quantitative comparisons. The experimental images are shown in Fig. 1(a), of which the compression strength is set as $Q = 60$. The MATLAB codes were run on a PC with 8G RAM and 2.6G Hz CPU.

As our task is to enhance the low lightness and suppress the JPEG artifact, we first use two baseline models for comparison. **Baseline 1**: remove the artifact at first and then enhance the low lightness. **Baseline 2**: enhance the low lightness at first and then

remove the artifact. For the fairness of the comparison, we use the methods mentioned in Sects. 3.3 and 3.4 to achieve the lightness enhancement and the artifact suppression, respectively. We also compare our method with the most related one proposed in [1] and term it as **ECCV14**. Its parameters are empirically set as the default values as in [1]. As for our methods, since the framework of our method is open for the choice of image enhancement models, we choose two kinds of them. The first is the one mentioned in Sect. 3.3, and we term it as **Ours-MF**. The second is the simple gamma correction process, and we term it as **Ours-GC**. Here we simply set the enhancement parameter γ as $1/2.2$ according to [8]. As for the image decomposition stage of **Ours-GC** and **Ours-MF**, we follow [8] to set $\eta_1 = 0.001$, $\eta_2 = 0.0001$, $\eta_3 = 0.25$, and $\varepsilon = 0.01$.

We first present the visual comparison in Fig. 4. From the two examples, we have the following observations. First, we observe the lost image details in the results of **Baseline 1** and **ECCV14**. The reason is that the image details hidden in the darkness are vulnerable to the JPEG removal process, and many of them are unnecessarily removed. In contrary, the artifact removal of our method is applied on the decomposed reflectance layer that extracts the image contents at various scales in advance. Second, **Baseline 2** well preserves the image details, but removes much fewer block artifacts than **Ours-MF** and **Ours-GC**. Since **Baseline 2** firstly enhances the input image, the originally weak block edges hidden in the darkness are unnecessarily boosted, which brings difficulty for the following suppression step. Our methods do not have this

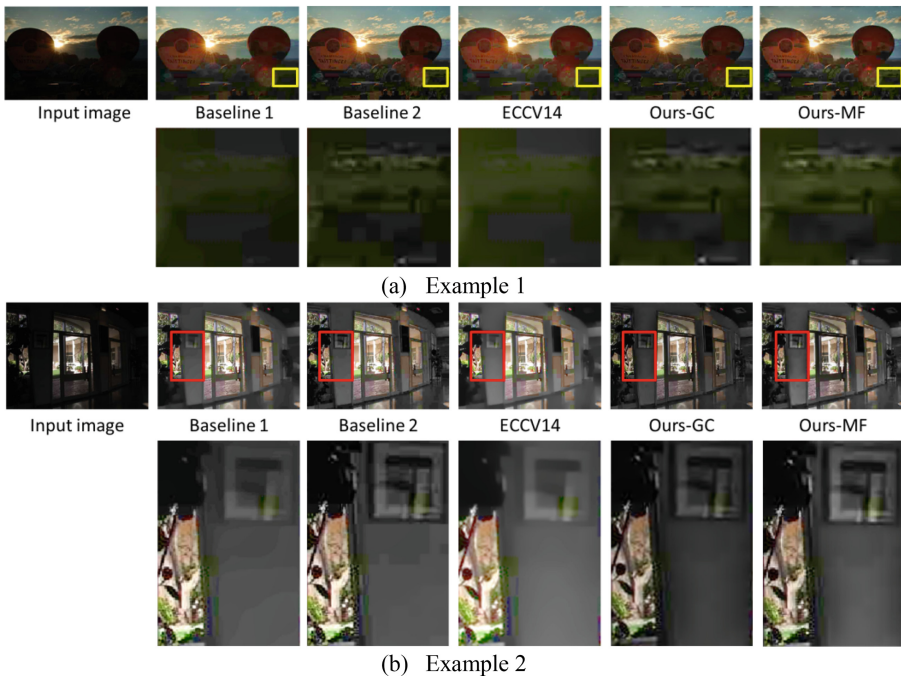


Fig. 4. Qualitative comparison between our methods and the counterparts.

problem due to the well-separated image layers. Third, by comparing the two versions based on our framework, we can see that **Ours-MF** achieves better results than **Ours-GC** in terms of the lightness condition.

We also quantitatively evaluate all the above methods with two metrics proposed in [13, 14]. The results are shown in Tables 1 and 2, in which the bold/underline numbers indicate the best/second best results among all the methods, respectively. From Table 1, we can see that **Ours-MF** and **Ours-GC** generally achieve better performance than other three methods in terms of removing JPEG artifacts. From Table 2, **Ours-MF** generally has the best performance in terms of image contrast. Differently, the performance of **Ours-GC** is less competitive than **ECCV14**. The reason is that the gamma correction model only imposes a global non-linear transform on the lightness layer. Based on the qualitative and quantitative comparison, **Ours-MF** achieves the best performance, and validates the effectiveness of our idea of decoupling the low lightness issue and the JPEG artifact issue at the beginning. In a word, all the above results validate the effectiveness of each part of our proposed framework.

Table 1. Quantitative comparison based on the metric measuring the block effect [13]

	Input	Baseline1	Baseline2	ECCV14	Ours-MF	Ours-GC
1	0.2109	0.0383	0.0296	<u>0.0276</u>	0.0119	0.0366
2	0.1975	0.0307	0.0216	0.0165	0.0037	<u>0.0045</u>
3	0.1900	0.0735	0.0472	0.0180	0.0064	<u>0.0074</u>
4	0.1941	0.0503	0.0407	0.0232	<u>0.0150</u>	0.0133
5	0.1894	0.0284	0.0171	0.0110	0.0021	<u>0.0034</u>
6	0.1268	<u>0.0445</u>	0.0195	0.0171	<u>0.0043</u>	0.0038
Average	0.1848	0.0443	0.0293	0.0189	0.0072	<u>0.0115</u>

The bold font means the best result

Table 2. Quantitative comparison based on the metric measuring the image contrast [14]

	Input	Baseline1	Baseline2	ECCV14	Ours-MF	Ours-GC
1	7.6885	7.1351	7.0683	<u>6.0436</u>	5.8170	7.6696
2	6.3567	4.7280	<u>4.0690</u>	4.6849	3.3641	4.3921
3	7.1700	5.3822	3.8854	<u>3.5243</u>	3.1429	3.7367
4	6.3027	5.0812	4.8454	<u>4.3957</u>	4.2165	4.9111
5	6.4311	2.9499	<u>2.9322</u>	2.5457	3.1349	3.1649
6	7.2752	<u>3.8666</u>	3.9664	3.5832	4.1252	5.3496
Average	6.8707	4.8571	4.4611	<u>4.1296</u>	3.9668	4.8707

The underline font means the second best result

5 Conclusions

In this paper, we propose an image enhancement method for compressed low-light images. Based on image decomposition, low lightness and JPEG artifact are separated into two decomposed layers, which can be well addressed by off-the-shelf enhancing

techniques. Visual and quantitative comparisons demonstrate the effectiveness of our method. We plan to improve our method by introducing saliency detection [15, 16].

Acknowledgements. The research was supported by the National Nature Science Foundation of China (Grant Nos. 61632007, 61772171, and 61702156), Anhui Provincial Natural Science Foundation (Grant No. 1808085QF188)

References

1. Li, Y., Guo, F., Tan, R.T., Brown, M.S.: A contrast enhancement framework with JPEG artifacts suppression. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds.) ECCV 2014. LNCS, vol. 8690, pp. 174–188. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-10605-2_12
2. Reza, A.M.: Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement. *J. VLSI Signal Process. Syst.* **38**(1), 35–44 (2004)
3. Yue, H., Yang, J., Sun, X., Wu, F., Hou, C.: Contrast enhancement based on intrinsic image decomposition. *IEEE TIP* **26**(8), 3981–3994 (2017)
4. Li, M., Liu, J., Yang, W., Sun, X., Guo, Z.: Structure-revealing low-light image enhancement via robust Retinex model. *IEEE TIP* **27**(6), 2828–2841 (2018)
5. Guo, X., Li, Y., Ling, H.: LIME: low-light image enhancement via illumination map estimation. *IEEE TIP* **26**(2), 982–993 (2017)
6. Lee, C., Lee, C., Kim, C.: Contrast enhancement based on layered difference representation of 2D histograms. *IEEE TIP* **22**(12), 5372–5384 (2013)
7. Fu, X., Zeng, D., Huang, Y., Zhang, X., Ding, X.: A weighted variational model for simultaneous reflectance and illumination estimation. In: Proceedings of CVPR (2016)
8. Cai, B., Xu, X., Guo, K., Jia, K., Hu, B., Tao, D.: A joint intrinsic-extrinsic prior model for retinex. In: Proceedings of ICCV (2017)
9. Kou, F., Wei, Z., Chen, W., Wu, X., Wen, C., Li, Z.: Intelligent detail enhancement for exposure fusion. *IEEE TMM* **20**(2), 484–495 (2018)
10. Ma, K., Li, H., Yong, H., Wang, Z., Meng, D., Zhang, L.: Robust multi-exposure image fusion: a structural patch decomposition approach. *IEEE TIP* **26**(5), 2519–2532 (2017)
11. Ying, Z., Li, G., Gao, W.: A bio-inspired multi-exposure fusion framework for low-light image enhancement. ArXiv, abs/1711.00591 (2017)
12. Fu, X., Zeng, D., Huang, Y., Liao, Y., Ding, X., Paisley, J.: A fusion-based enhancing method for weakly illuminated images. *Sig. Process.* **129**, 82–96 (2016)
13. Golestaneh, S.A., Chandler, D.M.: No-reference quality assessment of JPEG images via a quality relevance map. *IEEE SPL* **21**(2), 155–158 (2014)
14. Gu, K., Wang, S., Zhai, G.: Blind quality assessment of tone-mapped images via analysis of information, naturalness, and structure. *IEEE TMM* **18**(3), 432–443 (2016)
15. Jian, M., Qi, Q., Dong, J., Sun, X., Sun, Y., Lam, K.: Saliency detection using quaternionic distance based weber local descriptor & level priors. *MTAP* **77**(11), 14343–14360 (2018)
16. Jian, M., Lam, K., Dong, Y., Shen, L.: Visual-patch-attention-aware saliency detection. *IEEE Trans. Cybernetics* **45**(8), 1575–1586 (2015)