PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

Low-light image enhancement based on variational Retinex model

Li, Xuesong, Zhang, Hongyi, Pan, Jinfeng, Li, Qilei, Zou, Guofeng, et al.

Xuesong Li, Hongyi Zhang, Jinfeng Pan, Qilei Li, Guofeng Zou, Mingliang Gao, "Low-light image enhancement based on variational Retinex model," Proc. SPIE 12171, Thirteenth International Conference on Signal Processing Systems (ICSPS 2021), 1217108 (1 May 2022); doi: 10.1117/12.2631428

Event: Thirteenth International Conference on Signal Processing Systems (ICSPS 2021), 2021, Shanghai, China

Low-light Image Enhancement Based on Variational Retinex Model

Xuesong Li[a], Hongyi Zhang [a], Jinfeng Pan*[a], Qilei Li[b], Guofeng Zou[a], Mingliang Gao*[a]
[a]School of Electrical and Electronic Engineering, Shandong University of Technology, Zibo 255000, China; [b]School of Electronic Engineering and Computer Science, Queen Mary University of London, London, E1 4NS, United Kingdom.

ABSTRACT

The low-light image enhancement plays a crucial role in computer vision and multimedia applications. However, it is still a challenging task, as the degraded images reduce the visual naturalness and visibility. To address this problem, we build a novel variational Retinex model to accurately estimate the illumination and reflectance components. The illumination and reflectance are jointly updated by alternating optimization algorithm. Experimental results on several public datasets demonstrate that the proposed method outperforms the state-of-the-art methods in Retinex decomposition and illumination adjustment.

Keywords: Low-light image enhancement, variational Retinex model, illumination adjustment

1. INTRODUCTION

Taken in low-light conditions such as darkness and nighttime, low-light images suffer from unpleasant visual aesthetics, enormous noise, low contrast and color distortions. These degraded images will affect the performance of subsequent computer vision algorithms. Besides, with the prevalence of portable imaging devices, the demand for high-quality images with clear details and satisfied brightness becomes extremely imperative. Therefore, the demand of effective yet robust model for low-light image enhancement under various realistic scenes is highly urgent. The methods of low-light enhancement can be generally divided into three categories, namely histogram equalization-based methods[1], Retinex decomposition-based methods[2], and deep learning-based methods[3]. The Histogram equalization (HE)-based methods enhance the visibility of low-light images by flattening the histogram via stretching the corresponding dynamic range of the intensity[4, 5]. HE-based methods can be further classified into global HE-based methods and local HE-based methods. Although these methods are effective for dynamic range image enhancement, the enhanced image often exhibits unnatural details.

The Retinex decomposition-based methods enhance low-light images by image decomposition. These methods decompose the images into two components, namely reflectance and illumination. Then, these two components are further processed to obtain the enhanced results. The single-scaled Retinex (SSR)[6] method and multi-scaled Retinex (MSR)[7] method are the pioneering works in this field. Sequential methods consider both the illumination and reflectance layers to improve the performance[8]. However, it's inherently an ill-posed problem to estimate illumination and reflectance components from a single image. In order to make the problem trackable, some attempts transform the illumination or reflectance decomposition into a statistical reasoning problem and seek the optimal solutions by proposing different priors for illumination and reflectance and defining variational optimization[9].

The learning-based methods model the feature maps from the high visual quality images to enhance the low-light images. Lore[3] *et al*. first enhanced the low-light images by the stacking sparse auto-encoders. Subsequently, varied networks and diversified losses were proposed[10]. Besides, adversarial learning was introduced to obtain visual attributes beyond traditional metrics[11, 12]. Jiang[11] *et al*. proposed an EnlightenGAN to get rid of the construction of pairwise datasets. Although the deep learning- based approaches have achieved remarkable achievements in the domain of low-light image enhancement, the enormous computational burden in practical application and the complex structure of the model limits their popularity on mobile devices. Moreover, the learning-based methods rely heavily on great deal of high-quality images.

* Jinfeng Pan: pjfbysj@163.com; * Mingliang Gao: sdut_mlgao@163.com

Thirteenth International Conference on Signal Processing Systems (ICSPS 2021),
edited by Qingli Li, Kezhi Mao, Yi Xie, Proc. of SPIE Vol. 12171, 1217108
© 2022 SPIE · 0277-786X · doi: 10.1117/12.2631428

In this paper, we conduct an effective method to accomplish the structure and texture estimation during the Retinex decomposition. The proposed model is based on two highly correlated hypotheses, *i.e.,* the illumination component should be piece-wise smooth, while the reflectance component should contain as much detail as possible. Based on the structure and texture diagrams, we build a variational Retinex model to accurately estimate the illumination and reflectance components. Experiments on several challenging benchmarks verify the effectiveness of the proposed model in both subjective and objective evaluations.

2. METHODOLOGY

2.1 Retinex Model

Retinex theory[13] postulates that the input low-light image $O \in \mathbb{R}^{n \times m}$ can be represented as the product of illumination. $L \in \mathbb{R}^{n \times m}$ and reflectance $R \in \mathbb{R}^{n \times m}$ as $O = L \odot R$. The symbol $\odot$ means the element-wise multiplication. The decomposed components can be converted back by estimating them alternatively by $L = O \oslash R$ and $L = O \oslash R$, where $\oslash$ represents the element-wise division. Retinex theory introduces a valuable derivative property[14], *i.e.,* the larger derivative value in the image is usually due to the variation of the reflectance component, while the smaller derivative value may be due to the smooth distribution of the illumination. According to the properties of the image in the gradient field, previous variational Retinex methods generally utilize a variational objective function to estimate the illumination and reflectance components[15]. The objective function is formulated as

$$\min_{L,R} \| O - L \odot R \|_F^2 + \mathcal{N}_1(L) + \mathcal{N}_2(R) , \tag{1}$$

where $\mathcal{N}_1$ and $\mathcal{N}_2$ are regularization terms.

2.2 Structure and Texture Estimator

The regularization terms $\mathcal{N}_1$ and $\mathcal{N}_2$ in Eq. (1) can extract the structure and texture maps by distinguishing the difference in the distribution of gradients between the illumination and reflectance components. The smaller gradient is due to the smooth in illumination[14], while the larger gradient is due to the variations in reflectance. These local derivatives can reflect the content structure or texture through exponential growth or decay. For this purpose, we adopt a novel relative total variation to construct the structure and the texture prior matrix. The structure prior is to enforce the spatial smooth on illumination layer while preserve the main structure in illumination layer.

The formulation of the structure prior is given as

$$\mathcal{S}(O_p) = \sum_{q \in \mathcal{R}_p} \left(\frac{\mathcal{G}_\sigma * |\triangledown_x O_q|}{|\mathcal{G}_\sigma * \triangledown_x O_q| + \epsilon} + \frac{\mathcal{G}_\sigma * |\triangledown_y O_q|}{|\mathcal{G}_\sigma * \triangledown_y O_q| + \epsilon} \right)^{\gamma_s} \tag{2}$$

The proposed texture prior is to enforce the reflectance component to be piece-wise continuous. The texture prior is formulated as

$$\mathcal{T}(O_p) = \sum_{q \in \mathcal{R}_p} \frac{1}{\left(\left| \frac{\mathcal{G}_\sigma * |\triangledown_x O_q|}{|\mathcal{G}_\sigma * \triangledown_x O_q| + \epsilon} + \frac{\mathcal{G}_\sigma * |\triangledown_y O_q|}{|\mathcal{G}_\sigma * \triangledown_y O_q| + \epsilon} \right| + \varepsilon \right)^{\gamma_t}} \tag{3}$$

where ϵ =0.001 and ε =0.005. γ_s and γ_t are the structure and texture perception coefficients. Where $\triangledown_{x/y}$ is partial derivative in horizontal or vertical directions. The symbol $*$ is a convolutional operator. $\mathcal{G}_\sigma$ is the standard Gaussian kernel with window size σ = 3. $\mathcal{R}_p$ is a rectangular region centered on the pixel *p*, and *q* is the pixel belongs to $\mathcal{R}_p$.

2.3 Proposed Model

The proposed model is formulated as

$$argmin_{L,R} \|O - L \odot R\|_F^2 + \alpha \|S \odot \triangledown L\|_F^2 + \beta \|T \odot \triangledown R\|_F^2 + \lambda \|L - B\|_F^2 , \tag{4}$$

where α , β and λ control the importance of different terms in the objective function. $\|O - L \odot R\|_F^2$ constrains the fidelity between the observed image O and the reconstructed image $L \odot R$. $\|\mathcal{S} \odot \nabla L\|_F^2$ and $\|\mathcal{T} \odot \nabla R\|_F^2$ correspond the structure map of illumination component and texture map of texture map. $\|L - B\|_F^2$ minimizes the distance between the estimated illumination L and the initial illumination B .
The formulations of the second and third terms in Eq. (4) are denoted as

$$\|\mathcal{S} \odot \nabla L\|_F^2 = s_x \|\nabla_x L\|_F^2 + s_y \|\nabla_y L\|_F^2 , \tag{5}$$

$$\|\mathcal{T} \odot \nabla R\|_F^2 = t_x \|\nabla_x R\|_F^2 + t_y \|\nabla_y R\|_F^2 , \tag{6}$$

where

$$s_{x/y} = \left(\frac{\mathcal{G}_\sigma * |\nabla_x L|}{|\mathcal{G}_\sigma * \nabla_x L| + \epsilon}\right)^{\gamma_s} , \tag{7}$$

$$t_{x/y} = \frac{1}{\left(\left|\frac{\mathcal{G}_\sigma * |\nabla_x R|}{|\mathcal{G}_\sigma * \nabla_x R| + \epsilon} + \right| + \varepsilon\right)^{\gamma_t}} . \tag{8}$$

2.4 Optimization Algorithm

Denote that L_k and R_k are the illumination and reflectance components at the *k*-th iteration (*k*=0, 1, 2, ..., *K*), respectively. *K* is the maximum number of iterations. Two separated sub-problems are iteratively cycled through. We illuminate the solutions of the *k*-th iteration to the sub-problems as follows.

1) ***L* Sub-problem**. Neglecting the terms unrelated to *L* and initializing L_0=*O*, the optimization problem to *L* is formulated as

$$L_{k+1} = argmin_L \|O - L \odot R_k\|_F^2 + \alpha(s_x \|\nabla_x L\|_F^2 + s_y \|\nabla_y L\|_F^2) + \lambda \|L - B\|_F^2 . \tag{9}$$

To solve Eq. (9), the loss function to the matrix notation form is rewritten as

$$L_{k+1} = (L \odot R_k - O)^T (L \odot R_k - O) + \alpha(L^T D_x^T S_x D_x L + L^T D_y^T S_y D_y L) + \lambda(L - B)^T (L - B), \tag{10}$$

where D_x and D_y are the Toeplitz matrices in horizontal and vertical directions. S_x=*diag*(s_x) and S_y=*diag*(s_y). Then, the solution to Eq. (9) is

$$L_{k+1} = \frac{R_k^T O + \lambda B}{R_k^T R_k + a(D_x^T S_x D_x + D_y^T S_y D_y) + \lambda 1} \tag{11}$$

where ***1*** is an identity matrix.

2) ***R* Sub-problem.** We initialize the $R_0 = O \oslash L_1$ and update *R* while fixing *L*. The terms unrelated to *R* are neglected, and the following optimization problem is derived.

$$R_{k+1} = argmin_R \|O - L_{k+1} \odot R\|_F^2 + \beta(t_x \|\nabla_x R\|_F^2 + t_y \|\nabla_y R\|_F^2). \tag{12}$$

The solution to Eq. (12) is similar to Eq. (12). Rewrite the loss function to the matrix notation norm and get the solution to Eq. (13):

$$R_{k+1} = \frac{L_{k+1}^T O}{L_{k+1}^T L_{k+1} + \beta(D_x^T T_x D_x + D_y^T T_y D_y)} . \tag{13}$$

The cycled optimization continues until the convergence conditions[16] are satisfied, or the iterations reaches a pre-defined threshold. The summary of the optimization method for the proposed model is demonstrated in Figure. 1.

Algorithm 1: The optimization of the proposed model

Input: Observed image O, parameters γ_s, γ_t, α, β and λ, maximum iterations K and stopping parameters δ.

Initializing L_0 and R_0, and setting the structure and texture weight matrices $\mathcal{S}_0$ and $\mathcal{T}_0$.
for $k = 1 : K$ **do**
  1. Compute structure weight $\mathcal{S}_{k+1}$ by Eq. (7)
  2. Update L_{k+1} by Eq. (11)
  3. Compute texture weight $\mathcal{T}_{k+1}$ by Eq. (8)
  4. Update R_{k+1} by Eq. (13)
  if $\|L_{k+1} - L_k\|_F \,/\, \|L_k\|_F \leq \delta$ *or* $\|R_{k+1} - R_k\|_F \,/\, \|R_k\|_F \leq \delta$ **then**
    Stop Updating
    else
      Continue
    end
  end
end

Output: Estimated illumination and reflectance components.

Figure 1. The summary of the optimization method for the proposed model

2.5 Illumination Adjustment

After obtaining the enhanced components of illumination L, we adopt Gamma correction[17,18], $\hat{L} = L^{\frac{1}{\gamma}}$, to adjust the illumination component. Then, the enhanced result $\hat{O}$ is generated by $\hat{O} = R \odot L^{\frac{1}{\gamma}}$, where γ is empirically set to 2.2[19,20]. Finally, the enhanced image is reversed from HSV to RGB to get the final result.

3. EXPERIMENTAL RESULTS AND ANALYSIS

3.1 Experiment Settings and Implementation Details

The experiments are performed on a PC with an Intel i5-10400 CPU, 2.90GHz and 16GB memory. We set the parameters as γ_s=1.0, γ_t=0.75, α=0.001, β=0.0001, δ=0.005, and λ=0.25. For a fair comparison, the results of the compared methods come either from the original papers or reproduced by the official codes. The proposed is compared with 6 SOTA methods, including Dong[21], CVC[22], SSR[6], NPE[23], MF[24], and LIME[25]. The comparisons are conducted on 6 benchmarks, *i.e.*, LIME[25], DICM[26], MEF[27], NPE[23], LOL[28], LOE[29].

3.2 Retinex Decomposition

The performance of the proposed model on image decomposition is depicted in Figure. 2. The top row is the illumination component and the bottom row is the reflectance component. It shows that the illumination layer decomposed by SSR[6] in Figure. 2 (a) is suffered from serious pixel dislocation, which will introduce the "ghost effect". LIME[25] in Figure. 2 (b) can obtain reasonable illumination map, but the reflectance fails to capture the details. Comparatively speaking, the proposed model can enhance the spatial piece-wise smoothness while preserving the structural information. Also, the image details can be faithfully preserved across the whole reflectance component.

3.3 Qualitative Evaluation

The qualitative results are depicted in Figure. 3. Some observations can be derived as follows. First, the HE-based method, *e.g.*, CVC[22], tends to preserve the details in image but cannot improve the brightness effectively. For example, in Figure. 3 (c), the brightness is barely enhanced compared with the original image. Second, the result based on the SSR[6] has serious artifacts, *e.g.*, unrealistic edges, strongly boosted noise, and color distortion, as shown in Figure. 3 (d). Third, the methods of LIME[25] and Dong are effective in improving the image brightness, but they are suffered from artifacts such as over-enhancement or noise amplification. For instance, in Figure. 3 (b) and Figure. 3 (e), the edges of plant leaves are slightly blurred. Comparatively speaking, the NPE[23] and our method achieve satisfying visual quality in all the tested images, but our results are relatively more natural.

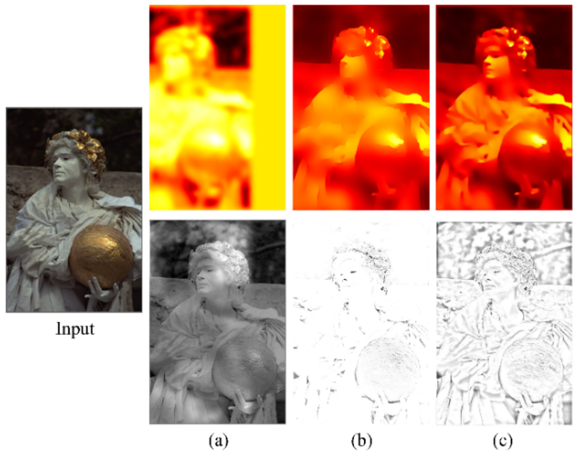

Figure 2. Decomposition results of the image "Statue". The first row are the estimated illumination maps and the second row are the estimated reflectance maps. (a) SSR[6]. (b) LIME[25] (c) Ours.

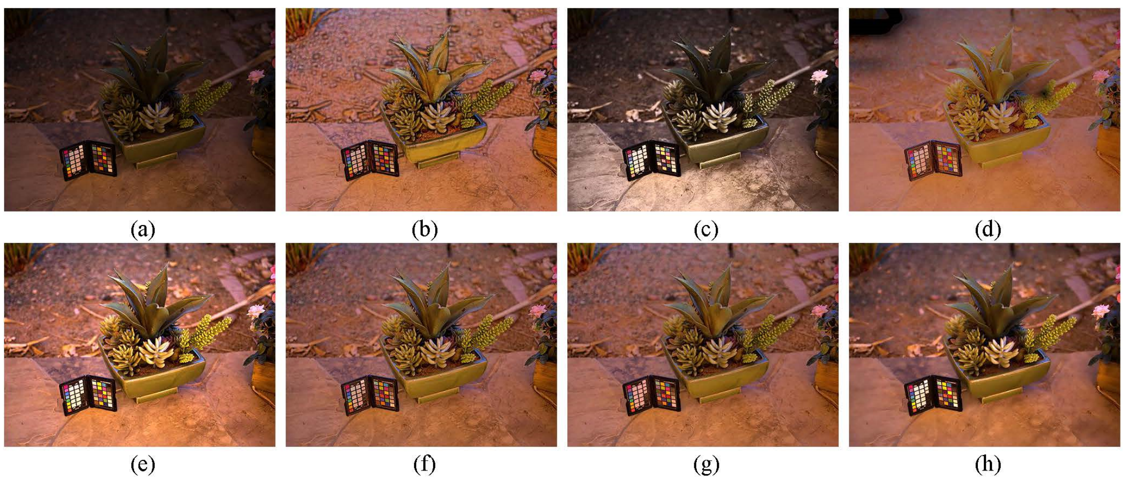

Figure 3. Qualitative evaluation with different methods. (a) Input (b) Dong[21] (c) CVC[22] (d) LDR[31] (e) LIME[25] (f) MF[24] (g) NPE[23] (h) Ours.

Table 1. Quantitative comparisons in terms of NIQE. The best score is denoted in bold.

Datasets / Methods	LIME[25]	DICM[26]	NPE[23]	LOE[29]	LOL[28]	MEF[27]	Average
Dong[21]	4.3240	3.7747	4.2652	4.4903	3.8842	3.4191	4.0263
CVC[22]	3.7775	**2.8245**	**3.1351**	4.7309	5.1495	2.8315	3.7415
SSR[6]	4.2358	3.4375	3.1238	4.4214	4.1085	2.7745	3.6836
NPE[23]	4.0751	3.2256	3.2540	**4.3179**	4.3507	**2.7394**	3.6605
MF[24]	4.1301	3.1081	3.3805	4.8837	4.2423	2.7499	3.7491
LIME[25]	4.5209	3.2488	3.4825	4.7246	4.1112	2.8096	3.8162
Ours	**3.6212**	3.0632	3.3322	4.6490	**3.3730**	3.6340	**3.6121**

3.4 Quantitative Evaluation

Considering that there is rare ground truth image in the dataset, we employ the NIQE indicator for quantitative evaluation. The NIQE gains the feature distribution via analyzing the dataset with high quality images. Then, it assesses the quality of an image by comparing the difference between the input image's feature distribution[32] with the sample feature distribution. A lower NIQE value indicates higher quality of an image. Quantitative results in terms of NIQE are shown in Table 1. It shows that the proposed method performs best in LIME[25] and LOL[28] datasets. Meanwhile, it achieves an average score of 3.6121 which outperforms the SOTA methods.

4. CONCLUSION

In this paper, we propose a novel variational Retinex model for low-light image enhancement. We explore the priors of structure and texture on illumination and reflectance components. The key idea is to accurately estimate the structure and texture maps via analyzing the difference of gradient distribution in illumination and reflectance layers. The proposed model is solved by an alternative update algorithm. To demonstrate the effectiveness of the proposed model, experiments are carried on six public datasets. The subjective and objective evaluations demonstrate that the proposed method outperforms the state-of-the-art methods in both Retinex decomposition and illumination adjustment.

ACKNOWLEDGMENTS

This work is supported by the National Natural Science Foundation of China (No.61801272), and National Natural Science Foundation of Shandong (No. ZR2020MF127).

REFERENCES

[1] Ibrahim, H. and Pik Kong, N.S., “Brightness preserving dynamic histogram equalization for image contrast enhancement,” IEEE Transactions on Consumer Electronics53(4), 1752–1758 (2007).

[2] Lee, C.-H., Shih, J.-L., Lien, C., and Han, C.-C., “Adaptive multiscale retinex for image contrast en-hancement,”2013 International Conference on Signal-Image Technology & Internet-Based Systems, 43–50(2013).

[3] Lore, K. G., Akintayo, A., and Sarkar, S., “Llnet: A deep autoencoder approach to natural low-light image enhancement,” Pattern Recognition61, 650–662 (2017)

[4] Singh, R. P. and Dixit, M., “Histogram equalization: a strong technique for image enhancement,” International Journal of Signal Processing, Image Processing and Pattern Recognition8(8), 345–352 (2015).

[5] Chin Jie Lew, Kok Swee Sim, and Teck Kiang Kho, "Contrast and Brightness Enhancement for DICOM Images and Lesions Auto Detection," Journal of Image and Graphics, Vol. 4, No. 1, pp. 25-28, June 2016. doi: 10.18178/joig.4.1.25-28

[6] Jobson, D., Rahman, Z., and Woodell, G., “Properties and performance of a center/surround retinex,” IEEE Transactions on Image Processing6(3), 451–462 (1997).

[7] Jobson, D., Rahman, Z., and Woodell, G., “A multiscale retinex for bridging the gap between color images and the human observation of scenes,” IEEE Transactions on Image Processing6(7), 965–976 (1997).

[8] Ren, X., Yang, W., Cheng, W.-H., and Liu, J., “Lr3m: Robust low-light enhancement via low-rank regularized retinex model,” IEEE Transactions on Image Processing29, 5862–5876 (2020).

[9] Fu, X., Liao, Y., Zeng, D., Huang, Y., Zhang, X., and Ding, X., “A probabilistic method for image enhancement with simultaneous illumination and reflectance estimation,” IEEE Transactions on Image Pro-cessing24, 4965–4977 (2015).

[10] Ma, L., Liu, R., Zhang, J., Fan, X., and Luo, Z., “Learning deep context-sensitive decomposition for low-light image enhancement,” IEEE Transactions on Neural Networks and Learning Systems, 1–15 (2021).

[11] Jiang, Y., Gong, X., Liu, D., Cheng, Y., Fang, C., Shen, X., Yang, J., Zhou, P., and Wang, Z., “Enlightengan:Deep light enhancement without paired supervision,” IEEE Transactions on Image Processing30, 2340–2349 (2021).

[12] Kim, G., Kwon, D., and Kwon, J., “Low-lightgan: Low-light enhancement via advanced generative adversarial network with task-driven training,” in [2019 IEEE International Conference on Image Processing (ICIP)], 2811–2815, IEEE (2019).

[13] Brainard, D. and Wandell, B., "Analysis of the retinex theory of color vision.," Journal of the Optical Society of America. A, Optics and image science3 10, 1651–61 (1986).
[14] Land, E., "The retinex theory of color vision," Scientific American237 6, 108–28 (1977).
[15] Liang, J. and Zhang, X., "Retinex by higher order total variationl1decomposition," Journal of Mathematical Imaging and Vision52, 345–355 (2015).
[16] Xu, J., Yu, M., Liu, L., Zhu, F., and Shao, L., "Star: A structure and texture aware retinex model," IEEE Transactions on Image Processing29, 5022–5037 (2020).
[17] Pang, J., Zhang, S., and Bai, W., "A novel framework for enhancement of the low lighting video," in [2017IEEE Symposium on Computers and Communications (ISCC)], 1366–1371 (2017).
[18] Wu, Y., Song, W., Zheng, J., and Liu, F., "Noisy low-light image enhancement using reflectance similarity prior," in [2020 15th IEEE International Conference on Signal Processing (ICSP)],1, 160–164 (2020).
[19] Gu, Z., Li, F., Fang, F., and Zhang, G., "A novel retinex-based fractional-order variational model for images with severely low light," IEEE Transactions on Image Processing29, 3239–3253 (2020).
[20] Gao, Y., Hu, H.-M., Li, B., and Guo, Q., "Naturalness preserved nonuniform illumination estimation for image enhancement based on retinex," IEEE Transactions on Multimedia20, 335–344 (2018).
[21] Dong, X., Pang, Y., and Wen, J., "Fast efficient algorithm for enhancement of low lighting video," Proceedings of the IEEE International Conference on Multimedia and Expo (ICME), 1–6 (2011).
[22] C ¸ elik, T. an Tjahjadi, T., "Contextual and variational contrast enhancement," IEEE Transactions on Image Processing20, 3431–3441 (2011).
[23] Wang, S., Zheng, J., Hu, H.-M., and Li, B., "Naturalness preserved enhancement algorithm for non-uniform illumination images," IEEE Transactions on Image Processing22, 3538–3548 (2013).
[24] Fu, X., Zeng, D., Huang, Y., Liao, Y., Ding, X., and Paisley, J., "A fusion-based enhancing method for weakly illuminated images," Signal Process.129, 82–96 (2016).
[25] Guo, X., Li, Y., and Ling, H., "Lime: Low-light image enhancement via illumination map estimation," IEEE Transactions on Image Processing26, 982–993 (2017).
[26] Lee, C. and Kim, C.-S., "Contrast enhancement based on layered difference representation," Proceedings of the IEEE International Conference on Image Processing (ICIP), 965–968 (2012).
[27] Ma, K., Duanmu, Z., Yeganeh, H., and Wang, Z., "Multi-exposure image fusion by optimizing a structural similarity index," IEEE Transactions on Computational Imaging4, 60–72 (2018).
[28] Wei, C., Wang, W., Yang, W., and Liu, J., "Deep retinex decomposition for low-light enhancement," in [Proceedings of the British Machine Vision Conference (BMVC)], (2018).
[29] Ying, Z., Li, G., and Gao, W., "A bio-inspired multi-exposure fusion framework for low-light image enhancement," ArXivabs/1711.00591(2017).
[30] Cai, B., Xu, X., Guo, K., Jia, K., Hu, B., and Tao, D., "A joint intrinsic-extrinsic prior model for retinex," Proceedings of the International Conference on Computer Vision (ICCV), 4020–4029 (2017).
[31] Lee, C. and Kim, C.-S., "Contrast enhancement based on layered difference representation of 2d histograms," IEEE Transactions on Image Processing22, 5372–5384 (2013).
[32] Mittal, A., Soundararajan, R., and Bovik, A., "Making a "completely blind" image quality analyzer," IEEE Signal Processing Letters20, 209–212 (2013).