ChinaRxiv

AI translation · View original & related papers at
chinarxiv.org/items/chinaxiv-202008.00091

Better-Than-Reference Low-Light Image Enhancement via Conditional Re-Enhancement Networks

Authors: Zhang Yu, Ti Xiaoguang, Zhang Bin, Ji Ruihang, Wang Chunhui, Wang Chunhui

Date: 2020-08-26T00:00:00+00:00

Abstract

Low light images suffer from severe noise, low brightness, low contrast, etc. In previous researches, many image enhancement methods have been proposed, but few methods can deal with these problems simultaneously. In this paper, to solve these problems simultaneously, we propose a low light image enhancement method that can combined with supervised learning and previous HSV (Hue, Saturation, Value) or Retinex model based image enhancement methods. First, we analyse the relationship between the HSV color space and the Retinex theory, and show that the V channel (V channel in HSV color space, equals the maximum channel in RGB color space) of the enhanced image can well represent the contrast and brightness enhancement process. Then, a data-driven conditional re-enhancement network (denoted as CRENet) is proposed. The network takes low light images as input and the enhanced V channel as condition, then it can re-enhance the contrast and brightness of the low light image and at the same time reduce noise and color distortion. It should be noted that during the training process, any paired images with different exposure time can be used for training, and there is no need to carefully select the supervised images which will save a lot. In addition, it takes less than 20 ms to process a color image with the resolution 400*600 on a 2080Ti GPU. Finally, some comparative experiments are implemented to prove the effectiveness of the method. The results show that the method proposed in this paper can significantly improve the quality of the enhanced image, and by combining with other image contrast enhancement methods, the final enhancement result can even be better than the reference image in contrast and brightness. (Code will be available at https://github.com/hitzhangyu/image-enhancement-with-denoise)

Full Text

Preamble

Better Than Reference In Low Light Image Enhancement: Conditional Re-Enhancement Networks

Yu Zhang, Xiaoguang Di, Bin Zhang, Ruihang Ji, and Chunhui Wang

Abstract—Low-light images suffer from severe noise, low brightness, and low contrast. While many image enhancement methods have been proposed in previous research, few can address these problems simultaneously. In this paper, we propose a low-light image enhancement method that combines supervised learning with previous HSV (Hue, Saturation, Value) or Retinex model-based image enhancement methods. First, we analyze the relationship between the HSV color space and Retinex theory, demonstrating that the V channel (the V channel in HSV color space, equivalent to the maximum channel in RGB color space) of the enhanced image can effectively represent the contrast and brightness enhancement process. Then, we propose a data-driven conditional re-enhancement network (denoted as CRENet). The network takes low-light images as input and uses the enhanced V channel as a condition to re-enhance the contrast and brightness of the low-light image while simultaneously reducing noise and color distortion. Notably, during training, any paired images with different exposure times can be used, eliminating the need for careful selection of supervised images and saving considerable effort. Additionally, it takes less than 20 ms to process a 400$×$600 color image on an RTX 2080Ti GPU. Finally, comparative experiments demonstrate the effectiveness of our method. The results show that the proposed method can significantly improve the quality of enhanced images, and when combined with other image contrast enhancement methods, the final enhancement result can even surpass the reference image in terms of contrast and brightness. (Code will be available at https://github.com/hitzhangyu/image-enhancement-with-denoise)

Index Terms—Low Light, Image Enhancement, Denoising, Color Correction

1 INTRODUCTION

When environmental lighting is low, such as at night or in a dark room, captured images become low-light images. These images always suffer from low contrast, low brightness, and serious noise. Low-light image enhancement methods are used to solve these problems before high-level tasks. In the past decades, researchers have proposed numerous non-learning-based image enhancement methods, such as [?], [?], [?], etc. Recently, with the development of deep learning, many supervised and unsupervised learning-based image enhancement methods have been proposed, such as [?], [?], [?], [?], etc., achieving promising results. However, whether these methods are based on learning or not, they have not been able to solve all these problems in low-light images well simultaneously.

For non-learning-based methods [?], [?], [?], most can significantly improve image contrast and brightness. However, it is difficult for these methods to reduce or suppress noise directly, and they may even amplify noise or cause color distortion during the enhancement process. Subsequent denoising operations often introduce problems such as blur and detail disappearance. For unsupervised image enhancement works, since it is difficult to introduce prior knowledge to the learning pipeline, especially regarding noise and color terms, there is always a problem of noise and color distortion [?]. For most supervised works [?], [?], there must be a hyperparameter to connect the input and reference images during training, caused by the non-one-to-one correspondence between input and reference images. This hyperparameter can adjust the contrast and brightness of the entire image during testing; however, it is hard to obtain the hyperparameter automatically, and the brightness and contrast of some local areas may remain unsatisfactory even with careful adjustment.

In this paper, we provide a low-light image enhancement framework that can integrate unsupervised learning or non-learning methods with supervised learning methods to simultaneously solve low contrast, low brightness, noise, and color distortion in low-light images.

The effectiveness of supervised methods often depends on data quality. However, in low-level image processing tasks, it is difficult to obtain input/label pairs like in high-level tasks, especially in image enhancement tasks. Although we can obtain images with different illumination from the same scene by changing lighting or modifying camera parameters, we cannot guarantee that the reference image has good contrast and brightness in every image block (e.g., Fig. 1 (e)-(f)). Meanwhile, since one low-light image can correspond to many high-light images, it is hard to ensure consistency in the selected data (similar input image blocks should correspond to similar reference image blocks in lighting). In fact, it can be considered that there is no ground-truth image here. The problem can be summarized as: how can we train networks without a unique or best ground-truth image, and how can we connect the input image and reference image when they do not meet consistency conditions during training? Different from previous supervised works that introduce a single hyperparameter such as time ratio [?] into the enhancement process, we propose a new framework that can use point-wise parameters related to contrast for training, and these parameters can be automatically obtained through other contrast enhancement methods during testing.

As shown in Fig. 1, combining our method with contrast enhancement methods like gamma correction allows us to adjust the contrast of the enhanced image and reduce noise simultaneously. Moreover, contrast in some local areas of the enhanced images is even better than in the reference images, such as bookcases and seats in dark areas of a stadium.

Similar to some previous works [?], we divide the low-light enhancement problem into two sub-problems: contrast and brightness enhancement, and denoising and color restoration. Different from previous methods that solve these sub-problems separately, we use contrast enhancement methods to generate point-wise contrast and brightness proposals and use them as a condition to re-enhance the low-light image while simultaneously reducing noise and color distortion. The method in this paper can be combined with any image contrast and brightness enhancement method based on HSV color space or Retinex model. The network can be trained without carefully selecting reference images, and any paired images with different lighting can be used for training.

Our contributions can be summarized as follows: * A conditional re-enhancement network (denoted as CRENet) for low-light image enhancement is proposed, which can simultaneously solve low contrast, low brightness, noise, and color distortion. CRENet can be combined with existing image enhancement methods and use the enhanced V channel as a condition to achieve re-enhancement of low-light images. In this process, CRENet can maintain the contrast and brightness of other enhancement methods while reducing noise and color distortion simultaneously. * Compared with other learning-based methods, the hyperparameters included in our method are point-wise and directly related to image contrast and brightness, making it possible for the enhanced image to have better contrast than the reference image by adjusting the brightness and contrast of local image areas—something difficult for other learning-based methods with only one hyperparameter. * By combining with other contrast and brightness enhancement methods, the proposed method does not need to carefully select reference images with good exposure, and any paired images with different brightness can be used for training.

2 RELATED WORKS

A. Non-learning based image enhancement methods

For non-learning-based single low-light image enhancement methods, there are mainly histogram equalization, methods based on dehazing or Retinex model, and other improved methods based on those approaches.

Histogram Equalization (HE) is one of the most widely used methods; however, it cannot avoid problems such as detail disappearance, poor color restoration, and noise amplification. Although many improved methods have been proposed to solve these problems [?], [?], [?], [?], [?], [?], many issues remain when applying histogram equalization directly to image enhancement.

Dong et al. [?] first proposed a method based on the dehazing model, and some studies extended these works [?]. Although these methods achieved some good effects, they lack a corresponding physical model, which limits the method' s application in various scenes.

Some works based on the Retinex model have been proposed to maintain image details and naturalness [?], [?], [?]; however, the denoising process before or after enhancement still causes blur or loss of details. To solve the noise and hole effect and preserve more details, many algorithms based on the variational

Retinex model have been proposed and achieved good results, such as [?], [?], [?], [?]; however, most of them cost too much time due to the need for multiple iterations to solve the variational equation.

Most non-learning methods focus on contrast and brightness enhancement, then use general denoising methods (like BM3D [?]) and white balance to remove noise and correct color; however, these methods always introduce blur and cannot solve the problem of color distortion. As shown in Fig. 2, although Gamma correction improves image contrast and brightness, after BM3D denoising, some details disappear.

B. Learning based image enhancement methods

As we know, LLNET [?] is the first work to use deep neural networks to solve image enhancement problems, proposing to train networks with synthetic noisy and dark images separately, but it does not consider the characteristics of natural images. Some other methods also use synthetic datasets, such as [?], [?], [?]. Although the data obtained by these methods appears to look like low-light images, it is difficult to truly reflect some characteristics of low-light images, such as noise, color distortion, and the coexistence of overexposed and underexposed areas in the same image.

Chen et al. [?] introduced a dataset containing real raw low-light images and corresponding raw high-light images for training, and they introduced an amplification ratio to connect the input and reference images, multiplying the ratio to a certain layer in the network. The ratio is set to the exposure time difference between the input and reference images during training. This method can solve the noise and color distortion problems well; however, the ratio must be chosen by the user during testing, which limits the method' s widespread use, and it cannot solve the inappropriate contrast problem well. It can be seen that this algorithm provides an exposure time adjustment method instead of physically adjusting exposure time. Obviously, only adjusting exposure time cannot work well in night scenes—there will still be some over-enhanced (saturated) areas or under-enhanced areas in the image. In our previous work [?], we provided a way to automatically learn the expected exposure time, but still could not solve the contrast problem because we can never provide ground-truth reference images with proper contrast in every local area, and without other constraints, it is difficult to solve this problem.

In [?], it was proposed to introduce the Retinex model into the training process to connect the reflection images of input and reference. However, due to the lack of constraints on the noise of the reflection image, additional denoising methods (like BM3D [?]) still need to be introduced. The method proposed in [?] resembles the combination of [?] and [?]. Compared to [?], [?] provides an extra brightness ratio to connect the illumination images of input and reference and adds a subnet called restoration-net to achieve denoising. However, during testing, it needs to manually adjust the ratio parameter to obtain better enhancement results. Although these methods use real low-light data for training, they do not constrain the contrast of the enhanced image, causing over-enhancement (saturation) or under-enhancement problems in the enhanced image.

In our previous work [?], we provided a max entropy Retinex model to achieve self-supervised learning while constraining contrast during training. However, due to the lack of strong constraints on color and noise, the enhanced image still looks like a nighttime image in color, and the noise cannot be removed well. During testing, we cannot guarantee good contrast in every local area.

Xiong et al. [?] decomposed the low-light image enhancement task into two stages: contrast enhancement and noise removal, proposing an unsupervised framework. However, in the absence of constraints on image contrast, the contrast and brightness after image enhancement may still be unsatisfactory. In this paper, we also prove that the Retinex model used in their first stage cannot ensure color information close to real daytime images.

Although these deep learning-based methods have achieved good visual effects in low-light image enhancement, they still cannot simultaneously solve low contrast, low brightness, noise, and color distortion.

3.1 Relationship between HSV color space and Retinex model

Recently, many low-light image enhancement works are based on the following Retinex model:

$$F(x) = R(x) \circ I(x)$$

where F and I represent the captured image and the illumination image respectively, R represents the reflection or desired image, and $\circ$ represents element-wise multiplication.

Most works assume that the three color channels of the image have the same illumination to simplify the model [?], [?], and generally use the maximum value of the three color channels as the initial estimate of the illumination map [?]. In the following description, we refer to this simplified Retinex model as the Retinex model.

Through simple transformation, it can be proven that image enhancement methods based on this simplified Retinex model are equivalent to performing enhancement operations on the V channel in HSV color space while leaving the H and S channels unchanged.

The color image can be divided into three channels according to each pixel’ s value in RGB color space:

$$\begin{cases} L(x) = \max_{c \in \{R,G,B\}} F_c(x) \\ M(x) = \text{median}_{c \in \{R,G,B\}} F_c(x) \\ N(x) = \min_{c \in \{R,G,B\}} F_c(x) \end{cases}$$

where x represents an individual pixel. Before image enhancement, the captured image F in HSV color space can be expressed as follows:

$$\begin{cases} V_b(x) = L(x) \\ S_b(x) = \frac{L(x) - N(x)}{L(x)} \\ H_b(x) = c_1 + c_2 \frac{M(x) - N(x)}{L(x) - N(x)} \end{cases}$$

where c_1 can be 60, 120, 240, and c_2 can be ± 60. Their values depend on the three color channels of the image, and for one image, c_1 and c_2 are certain values at each pixel. V_b, S_b, and H_b represent the V, S, and H channels of the low-light image before enhancement, respectively, with subscript b representing "before."

Based on the Retinex model, the desired image R can be obtained by the following formula:

$$R(x) = \frac{F(x)}{I(x) + \varepsilon}$$

where ε is a very small constant to avoid zero denominator. For simplicity, we omit ε in the following formulas and directly use $I(x)$ to represent $I(x) + \varepsilon$.

According to Equation (6), the enhanced image R in HSV color space can be expressed as follows:

$$\begin{cases} V_a(x) = \frac{L(x)}{I(x)} = \frac{V(x)_{\text{before}}}{I(x)} \\ S_a(x) = \frac{L(x)/I(x) - N(x)/I(x)}{L(x)/I(x)} = S_b(x) \\ H_a(x) = c_1 + c_2 \frac{M(x)/I(x) - N(x)/I(x)}{L(x)/I(x) - N(x)/I(x)} = H_b(x) \end{cases}$$

where V_a, S_a, and H_a represent the V, S, and H channels of the low-light image after enhancement, respectively, with subscript a representing "after."

It can be seen that the H and S channels of the captured image and the enhanced image are the same, and the enhancement operation only works on the V channel. Therefore, the enhanced V channel can well represent the image enhancement operation for methods based on HSV or Retinex model. It is obvious that hue and saturation differ between nighttime and daytime images, as well as between low-light and high-light images. This is caused by the non-linearity of the camera response curve, and even different color channels have different response curves. Thus, methods based on the simplified Retinex model cannot ensure that the color of enhanced images looks like real images captured in daylight or high-light conditions.

Therefore, we provide a supervised image enhancement method called the Conditional Re-Enhancement Network (CRNet) which takes the enhanced V channel as an additional point-wise parameter to allow the network to focus on learning the H and S channels that change with the V channel. Simultaneously, with the enhanced V channel as an additional parameter, we can benefit from previous research [?], [?], [?] on image contrast and brightness enhancement, and supervised learning also demonstrates good performance in noise suppression. In this way, we can simultaneously solve the problems of low contrast, low brightness, noise, and color distortion in low-light images.

Next, we provide the details of the network and training, and further explain the reason and rationality of using the V channel instead of a single parameter.

3.2 Conditional re-enhancement network

According to the camera response curve, a pixel in the input and reference image can be expressed as follows:

$$\begin{cases} X_{cij} = G_c(E_i \Delta t_j) \\ Y_{cik} = G_c(E_i \Delta t_k) \end{cases}$$

where X and Y represent the input low-light image and the reference image respectively, c represents one of the color channels, G_c represents the camera response curve of channel c, i represents the pixel, and Δt_j and Δt_k represent the exposure times of the low-light image and high-light image respectively. E_i represents the irradiance. Obviously, different reference images may have different exposure times.

Most supervised works should include at least one time-related item α to connect the input and reference images. Then, most learning-based models can be described as follows:

$$\max p(\theta | f_\theta, Y, X, \alpha)$$

where f and θ represent the enhancement network and its parameters respectively, and p represents the probability. α represents the hyperparameter related to time difference, used to connect X with manually selected Y. Previous works adopted time ratio $\alpha = \Delta t_k / \Delta t_j$ [?], average brightness ratio $\alpha = \mathrm{mean}(Y/X)$ [?], etc. As mentioned before, there are two obvious problems in previous works. Firstly, it is difficult to assign a value to α automatically during the application phase, whether it is time or average brightness difference, because those values are not directly related to image contrast. Secondly, a single parameter can only enhance the overall brightness of the image and cannot guarantee good contrast in every local area. This can be illustrated through the following equations.

According to Bayes' rule, by calculating the negative logarithm of Equation (11), the training phase can be expressed as follows:

$$\max p(\theta|f_\theta, Y, X, \alpha(X,Y)) = \min \|f_\theta(X, \alpha(X,Y)) - Y\|$$

If we assume that Y is the optimal reference in a series of images with different exposure times and we can manually select the best result during testing without other prior losses, the best result can be expressed as follows:

$$\hat{Y} = f_\theta(X, \alpha(X,Y)) = Y$$

where $\hat{Y}$ represents the enhanced image. It can be seen that the best result of the network is hard to exceed the manually selected reference Y. In addition, it is obvious that in many low-light scenes, we can never obtain the optimal image as a ground-truth reference by adjusting exposure time (e.g., Fig. 1 (e)-(f)), which means we cannot obtain a satisfactory image by only adjusting a single parameter α.

In fact, without other priors, the end-to-end supervised method cannot solve the problem of low contrast in low-light images, or even the problem of low brightness, unless we manually adjust the reference images carefully or use multiple images to synthesize reference images. However, this also brings other problems, such as increased time cost and the inability of the adjusted image to guarantee brightness consistency (the same low-light image patches correspond to different reference images). To solve these problems, we propose CRENet to explicitly control the contrast and brightness of the enhanced image. CRENet takes the V channel enhanced by other image contrast enhancement methods as a point-wise condition and can re-enhance the contrast and brightness of low-light images according to this condition.

The V channel is the maximum of the three color channels at each pixel, so the V channel of the reference image can be expressed with the camera response curve as follows:

$$V_{ik} = G_V(E_i \Delta t_k)$$

We make the same reasonable assumption that the function G is monotonically increasing [**?**]. Therefore, for a fixed pixel i, there is a certain Δt_k given V_{ik}, which means the V channel is sufficient to connect the other two channels between input and reference during training since the three color channels have the same exposure time. As shown in Fig. 3, with exposure changes, the R, G, B values at one pixel can form a curve in 3D space related to the camera response curve. Our motivation is to let the network learn this curve, and then with a given V value, it can obtain the values of the other two channels (HSV and RGB are just different color representations, so we directly implement it in RGB space when designing and training the network).

Thus, the V channel can be treated as a point-wise parameter that allows us to achieve different levels of enhancement in different areas. Through Equations (2) to (7), it can be seen that the V channel can well represent the processing results of other contrast enhancement methods.

During training, we can take the V channel of the reference images as point-wise parameter α, which can be expressed as follows:

$$\alpha = \max_{c\in\{R,G,B\}} Y_c + n$$

where n represents Gaussian noise, introduced to avoid identity transformation and simultaneously simulate noise in the V channel during testing.

Meanwhile, we can perform brightness mapping on the V channel of low-light images according to the reference to get α, which can be expressed as follows:

$$\alpha_i = w \max_{c\in\{R,G,B\}} Y_c$$

where w is the average brightness ratio calculated on every local area Ω around pixel i, expressed as follows:

$$w = \frac{\sum_{x\in\Omega(i)} Y(x)}{\sum_{x\in\Omega(i)} X(x)}$$

During testing, α can be expressed as follows:

$$\alpha = \max_{c\in\{R,G,B\}} g(X_c)$$

where g represents any contrast and brightness enhancement methods in HSV color space or based on Retinex model, such as histogram equalization, LIME [**?**], and self-supervised methods [**?**], etc.

The training procedure is achieved by minimizing the loss between the enhanced image and the supervised image, expressed as follows:

$$L = \|\hat{Y} - Y\|_1 + \mathrm{SSIM}(\hat{Y}, Y)$$

where SSIM represents structural similarity measurement [**?**]. We have also tested other loss functions, such as perceptual loss, color loss expressed by outer product like [?], loss in H and S channels in HSV space, gradient loss [?], etc., but the effect of these loss functions on the results is not obvious on the LOL dataset, neither in visual effects nor quantitative metrics.

The architecture of our CRENet is shown in Table 1, and the training pipeline is shown in Fig. 4. It should be noted that our method does not have special requirements for the network architecture. The network can be simplified or more carefully designed to reduce running time or further improve processing effects. We have tried more complex networks and found that a more complex structure may bring better visual effects. However, this is not the focus of this paper, so we do not show relevant results in the experiment section.

4.1 Implementation Details

The LOL (Low Light) dataset [?], which contains 500 image pairs, is used for training and testing. 485 image pairs from the database are used for training, with image size 400\$$\times 600. Duringtraining, thebatchsizeissetto 48 andthepatchsizeissetto 48 \times$\$48. We use Adam stochastic optimization [?] to train the network with a learning rate of 0.001. Network training and testing are implemented on an Nvidia GTX 2080Ti GPU and Intel Core i9-9900K CPU, with code based on the TensorFlow framework.

4.2 Performance Evaluation

In this section, we show results when combining with previous methods, including methods under HSV color space such as LAHE (Local Area Histogram Equalization) and Gamma correction, and methods based on Retinex model such as LIME [?], self-supervised image enhancement [?], Retinex-Net [?], KinD [?], etc. We also compare our method with BM3D in terms of denoising effects. Three metrics are adopted for quantitative comparison: PSNR, SSIM, and NIQE [?]. NIQE is a non-reference image quality assessment method that evaluates image naturalness, where lower values indicate better quality. PSNR and SSIM are reference image quality assessment methods indicating noise level and structural similarity between result and reference, respectively. Note that SSIM and PSNR are mainly used to evaluate structure and noise, but brightness differences seriously impact these metrics. To better evaluate results generated by different methods, we first perform local brightness mapping on the enhanced image to exclude brightness differences' influence before using SSIM and PSNR, following Equation (16). The mapping operation acts on the V channel while keeping H and S channels unchanged.

Table 2 and Fig. 5 show experimental results on the LOL dataset when combined with different image enhancement methods. In Table 2, it can be seen that for most existing methods, our method can significantly reduce noise, improve structural similarity, and make results more natural. As shown in Fig. 5, taking images enhanced by other methods as conditions, CRENet can maintain similar contrast and brightness to those enhanced images while simultaneously reducing noise and color distortion. Also, CRENet does not introduce obvious blur during denoising, and even for the LAHE method which produces severe noise, the re-enhancement network works well (e.g., Fig. 5 (b) and (j)). For the KinD [?] method, our method does not achieve PSNR and SSIM improvement because there is less noise and color distortion on the LOL dataset after KinD [?].

However, KinD can be treated as a whole image exposure time adjustment method, meaning it cannot ensure proper contrast or brightness in every local area. As shown in Fig. 6, we tried different ratios such as 5 and 8 (the author suggested maximum is 5). It can be seen that the lower right corner area of the image cannot be enhanced well even when changing the ratio. In fact, we even tried ratio parameter 100 in experiments, but still could not obtain good results.

We also study the influence of different sources of α on training: from the reference image with noise, from the low-light image after mapping, and from a mixture of the two. Results are shown in Table 3 and Fig. 7. When condition α comes from the mixture of low-light and high-light images, the network shows better results on evaluation metrics. In Fig. 7, when α comes from the reference, there is still some low-frequency and impulse noise (Fig. 7 (b)). However, if α comes from low-light images with brightness mapping, there is less noise. This is because we only add Gaussian noise to reference images when obtaining α, which cannot simulate real noise well during testing. This also shows that real low-light images contain far more complex noise than Gaussian noise, making it hard to obtain promising results in real low-light image enhancement tasks using only synthetic low-light images. Training with real data also brings problems in previous works, such as needing parameter adjustment during testing and careful reference selection during training, and enhanced results are difficult to exceed the reference. However, the proposed method can well solve these problems.

We also tested output consistency when input images are the same scene with different brightness, as shown in Fig. 8. The enhanced results basically maintain brightness consistency across the same blocks.

5 CONCLUSION

In this paper, we propose a conditional re-enhancement network for low-light image enhancement to simultaneously solve low contrast, low brightness, noise, and color distortion. The network can be combined with any contrast enhancement method based on HSV color space or simplified Retinex model, using the enhanced V channel by other methods as conditions to achieve re-enhancement. Experimental results demonstrate the method' s effectiveness and advantages. However, there are still some shortcomings: the final enhancement result depends on the contrast enhancement method, and image saturation is reduced in some areas (e.g., Fig. 1 (l)). Future research will focus on applying this technology in long-term and nighttime localization to eliminate light interference on feature extraction, generate corresponding nighttime datasets, and use extremely low-light images or raw images for training.

REFERENCES

[1] D. J. Jobson, Z.-u. Rahman, and G. A. Woodell, "A multiscale retinex for bridging the gap between color images and the human observation of scenes," IEEE Transactions on Image processing, vol. 6, no. 7, pp. 965–976, 1997.

[2] R. Kimmel, M. Elad, D. Shaked, R. Keshet, and I. Sobel, "A variational framework for retinex," International Journal of computer vision, vol. 52, no. 1, pp. 7–23, 2003.

[3] X. Guo, Y. Li, and H. Ling, "Lime: Low-light image enhancement via illumination map estimation," IEEE Transactions on image processing, vol. 26, no. 2, pp. 982–993, 2016.

[4] C. Wei, W. Wang, W. Yang, and J. Liu, "Deep retinex decomposition for low-light enhancement," arXiv preprint arXiv:1808.04560,

[5] Y. Zhang, J. Zhang, and X. Guo, "Kindling the darkness: A practical low-light image enhancer," in Proceedings of the 27th ACM International Conference on Multimedia, 2019, pp. 1632–1640.

[6] Y. Zhang, X. Di, B. Zhang, and C. Wang, "Self-supervised image enhancement network: Training with low light images only," arXiv, pp. arXiv–2002, 2020.

[7] W. Xiong, D. Liu, X. Shen, C. Fang, and J. Luo, "Unsupervised real-world low-light image enhancement with decoupled networks," arXiv preprint arXiv:2005.02818, 2020.

[8] S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. ter Haar Romeny, J. B. Zimmerman, and K. Zuiderveld, "Adaptive histogram equalization and its variations," Computer vision, graphics, and image processing, vol. 39, no. 3, pp. 355–368, 1987.

[9] C. Lee, C. Lee, and C.-S. Kim, "Contrast enhancement based on layered difference representation of 2d histograms," IEEE transactions on image processing, vol. 22, no. 12, pp. 5372–5384, 2013.

[10] C. Chen, Q. Chen, J. Xu, and V. Koltun, "Learning to see in the dark," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 3291–3300.

[11] K. G. Lore, A. Akintayo, and S. Sarkar, "Llnet: A deep autoencoder approach to natural low-light image enhancement," Pattern Recognition, vol. 61, pp. 650–662, 2015.

[12] E. D. Pisano, S. Zong, B. M. Hemminger, M. DeLuca, R. E. Johnston, K. Muller, M. P. Braeuning, and S. M. Pizer, "Contrast limited adaptive histogram

equalization image processing to improve the detection of simulated spiculations in dense mammograms," Journal of Digital imaging, vol. 11, no. 4, p. 193, 1998.

[13] Y.-T. Kim, "Contrast enhancement using brightness preserving bi-histogram equalization," IEEE transactions on Consumer Electronics, vol. 43, no. 1, pp. 1–8, 1997.

[14] Y. Wang, Q. Chen, and B. Zhang, "Image enhancement based on equal area dualistic sub-image histogram equalization method," IEEE Transactions on Consumer Electronics, vol. 45, no. 1, pp. 68–75,

[15] T. Celik and T. Tjahjadi, "Contextual and variational contrast enhancement," IEEE Transactions on Image Processing, vol. 20, no. 12, pp. 3431–3441, 2011.

[16] X. Dong, G. Wang, Y. Pang, W. Li, J. Wen, W. Meng, and Y. Lu, "Fast efficient algorithm for enhancement of low lighting video," in 2011 IEEE International Conference on Multimedia and Expo. IEEE, 2011, pp. 1–6.

[17] L. Li, R. Wang, W. Wang, and W. Gao, "A low-light image enhancement method for both denoising and contrast enlarging," in 2015 IEEE International Conference on Image Processing (ICIP). IEEE, 2015, pp. 3730–3734.

[18] S. Wang, J. Zheng, H.-M. Hu, and B. Li, "Naturalness preserved enhancement algorithm for non-uniform illumination images," IEEE Transactions on Image Processing, vol. 22, no. 9, pp. 3538–3548,

[19] X. Fu, D. Zeng, Y. Huang, Y. Liao, X. Ding, and J. Paisley, "A fusion-based enhancing method for weakly illuminated images," Signal Processing, vol. 129, pp. 82–96, 2016.

[20] X. Fu, D. Zeng, Y. Huang, X. Ding, and X.-P. Zhang, "A variational framework for single low light image enhancement using bright channel prior," in 2013 IEEE Global Conference on Signal and Information Processing. IEEE, 2013, pp. 1085–1088.

[21] S. Park, S. Yu, B. Moon, S. Ko, and J. Paik, "Low-light image enhancement using variational optimization-based retinex model," IEEE Transactions on Consumer Electronics, vol. 63, no. 2, pp. 178–184, 2017.

[22] G. Fu, L. Duan, and C. Xiao, "A hybrid l2-lp variational model for single low-light image enhancement with bright channel prior," in 2019 IEEE International Conference on Image Processing (ICIP). IEEE, 2019, pp. 1925–1929.

[23] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-d transform-domain collaborative filtering," IEEE Transactions on image processing, vol. 16, no. 8, pp. 2080–2095, 2007.

[24] C. Li, J. Guo, F. Porikli, and Y. Pang, "Lightennet: A convolutional neural network for weakly illuminated image enhancement,"Pattern Recognition Letters, vol. 104, pp. 15–22, 2018.

[25] J. Yang, X. Jiang, C. Pan, and C.-L. Liu, "Enhancement of low light level images with coupled dictionary learning," in 2016 23rd International Conference on Pattern Recognition (ICPR). IEEE, 2016, pp. 751–756.

[26] L. Shen, Z. Yue, F. Feng, Q. Chen, S. Liu, and J. Ma, "Msr-net: Low-light image enhancement using deep convolutional network," arXiv preprint arXiv:1711.02488, 2017.

[27] Q. Fu, X. Di, and Y. Zhang, "Learning an adaptive model for extreme low-light raw image processing," arXiv preprint arXiv:2004.10447, 2020.

[28] P. E. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," in ACM SIGGRAPH 2008 classes, 2008, pp. 1–10.

[29] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," IEEE transactions on image processing, vol. 13, no. 4, pp. 600–612,

[30] C. Guo, C. Li, J. Guo, C. C. Loy, J. Hou, S. Kwong, and R. Cong, "Zero-reference deep curve estimation for low-light image enhancement,"in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 1780–1789.

[31] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv preprint arXiv:1412.6980, 2014.

[32] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a"completely blind" image quality analyzer," IEEE Signal Processing Letters, vol. 20, no. 3, pp. 209–212, 2012.

Yu Zhang was born in Hebei, China. He received the B.E. degree and M.S. degree in Control Science and Engineering from Harbin Institute of Technology, China, in 2016 and 2018, respectively. He is currently pursuing a Ph.D. degree in Electronic Science and Technology at the National Key Laboratory of Tunable Laser Technology, Harbin Institute of Technology. His research interests include image restoration and enhancement, and SLAM.

Xiaoguang Di was born in Heilongjiang, China. He received the M.S. and Ph.D. degrees in Navigation, Guidance and Control from Northwestern Polytechnical University, China, in 1999 and 2004, respectively. He is currently an associate professor with the Control and Simulation Center, Harbin Institute of Technology, where he teaches courses in digital image processing and computer vision. His current research interests include real-time image restoration and enhancement, 3D object detection and recognition, and SLAM. Prof. Di is a member of the China Simulation Federation and Chinese Society of Astronautics.

Bin Zhang was born in Gansu, China. He received the B.Sc. degree in Applied Physics and the M.Sc. degree in Physics from China University of Petroleum, in 2010 and 2013, respectively. He is currently pursuing a Ph.D. degree at

the National Key Laboratory of Tunable Laser Technology, Harbin Institute of Technology. His research interests include laser imaging and laser transmission.

Ruihang Ji (S' 16) received the B.S. degree in Automation Engineering from Harbin Institute of Technology, Harbin, China, in 2016, where he is currently pursuing the Ph.D. degree in Control Science and Engineering. He is currently an exchange Ph.D. student supported by the China Scholarship Council with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore. His current research interests include adaptive control, robust control, unmanned aerial vehicles, and computer vision.

Chunhui Wang received the B.S., M.S., and Ph.D. degrees from Harbin Institute of Technology in 1987, 1991, and 2005, respectively. He is currently a professor and Deputy Director of the Institute of Optoelectronic Technology, Harbin Institute of Technology. His current major research interests are laser remote sensing, lidar, and laser detection and recognition.

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv —Machine translation. Verify with original.