

Low-light Image Enhancement Via Multispectral Reconstruction

Lei Xin^{1,2}, Xiangcai Ma³, Hang Luo¹, Xinrong Hu¹, Kaida Xiao⁴, Jinxing Liang^{1,4,*}

¹Wuhan Textile University, Wuhan 430200, China

²Wuchang University of Technology, Wuhan 430223, China

³Shanghai Publishing and Printing College, Shanghai 200093, China

⁴University of Leeds, Leeds, LS2 9JT, UK

*Email: jxliang@wtu.edu.cn

Abstract Low-light images often fail to accurately capture color and texture, limiting their practical applications in imaging technology. The use of low-light image enhancement technology can effectively restore the color and texture information contained in the image. However, current low-light image enhancement directly calculates from low-light to normal-light images, ignoring the basic principles of imaging, and the image enhancement effect is limited. The Retinex model splits the image into illumination components and reflection components, and uses the decomposed illumination components and reflection components to achieve end-to-end enhancement of low-light images. Inspired by the Retinex theory, this study proposes a low-light image enhancement method based on multispectral reconstruction, which first uses a multispectral reconstruction algorithm to reconstruct a metameric multispectral image of a normal-light RGB image, then uses a deep learning network to learn the end-to-end mapping relationship from a low-light RGB image to a normal-light multispectral image. In this way, any low-light image can be reconstructed into a normal-light multispectral image. Finally, the corresponding normal-light RGB image is calculated according to the colorimetry theory. To test the proposed method, the popular dataset for low-light image enhancement of LOL is adopted to compare the proposed method and the existing methods. During the test, a multispectral reconstruction method based on reversing the ISP of RGB imaging is used to reconstruct the corresponding metameric multispectral image of each normal-light RGB image in LOL, and the deep learning architecture proposed by Zhang *et al.* with CBAM added is used to establish the mapping relationship between the low-light RGB images and the corresponding reconstructed multispectral images. The proposed method is compared to existing methods such as Self-supervised, RetinexNet, RRM, KinD, RUAS, and Uretinex-net. In the context of the LOL dataset and an illuminant chosen for rendering, the results show that the low-light image enhancement method proposed in this study is better than the existing methods.

Index Terms: Color imaging; Low-light image enhancement; Imaging model; CBAM; Multispectral reconstruction

Introduction

Images taken with insufficient lighting intensity have defects such as underexposure and missing image texture details, which limit the extraction of image features and applications in the visual tasks. To address this problem, researchers have proposed many low-light image enhancement methods^[1-8]. The Retinex theory assumes that the image can be decomposed into two components of illumination and reflectance^[1], which is widely used in low-light enhancement tasks. Based on this theory, the Single-scale Retinex^[2], Multi-scale Retinex^[3]

were developed. Wang *et al.* proposed a method to enhance image details while maintaining lighting naturalness^[4]. Li *et al.* proposed a robust Retinex method by adding a noise term to handle low-light image enhancement under strong noise conditions^[5].

With the development of deep learning in the field of computer vision, neural networks have also been used in low-light enhancement. Enhancement networks based on deep learning usually use a large number of data sample pairs to extract data features and cooperate with nonlinear function fitting to obtain expected results. The KinD network, trained with image pairs taken under varying exposure conditions, can decompose images into reflection and illumination components^[6]. Benefiting from the progress of convolutional neural networks, Li *et al.* proposed LightenNet to enhance low-light images^[7]. Inspired by information entropy theory and the Retinex model, Zhang *et al.* proposed a self-supervised low-light image enhancement network, which only requires minutes of training to achieve image enhancement^[8]. In addition, Liu *et al.* proposed the RUAS method to construct a lightweight yet effective enhancement network for low-light images in real-world scenarios^[9]. Wu *et al.* proposed a Retinex-based deep unfolding network (URetinex-Net), which unfolds an optimization problem into a learnable network to decompose a low-light image into reflectance and illumination layers^[10]. Furthermore, some physical modeling methods and those integrate denoising and color restoration have also been studied recently^[11,12].

The above image enhancement methods overlook the camera imaging model which includes the light source, the object, and the camera. As we all know, when the light is incident on the surface of an object, the object will reflect part of the wavelength. The reflected radiance is acquired by the camera and forms an RGB image that satisfies human eye perception^[13]. In the field of color science, spectral reflectance is considered the fingerprint of color information. A multispectral image contains not only two-dimensional feature information but also three-dimensional spectral information. Among existing multispectral image acquisition methods, the multispectral reconstruction technology based on machine learning algorithms has received widespread attention from researchers. Finding the connection between RGB images and multispectral images through mathematical algorithms can provide fairly accurate multispectral approximations^[14].

Based on the strong correlation between multispectral images and RGB images, Zhang *et al.* proposed a deep-learning network on the basis of dense connections^[15]. The network first extracts shallow information from the RGB image then uses a 7-layer dense network to learn the deep mapping relationship of them, and finally outputs the multispectral image corresponding to the RGB image through a reconstruction layer including three layers of convolution. Li *et al.* proposed an adaptive weighted network based on an attention mechanism (Adaptive Weighted Attention Network, AWAN)^[16], it mainly consists of dual residual attention blocks (Dual Residual Attention Blocks, DRAB)^[17]. In order to achieve spectral reconstruction under unknown spectral response, Shi *et al.* proposed the HSCNN+ network based on the HSCNN network by using a convolutional layer instead of the up-sampling step^[18]. Then the residual network commonly used in computer image research is added to increase the network depth and improve reconstruction accuracy^[19]. After using the dense networks, a better spectral reconstruction accuracy was achieved^[20].

The Retinex theory, akin to the camera's perception process, equates the image's reflection component with the spectral reflectance in color visual perception, offering a theoretical foundation for low-light image enhancement based on spectral components. In the proposed method, we use a multispectral reconstruction method based on reversing the ISP of RGB imaging to obtain metameric multispectral images of normal-light RGB images, and apply the dense neural network proposed by Zhang *et al.* to establish the mapping relationship between low-light RGB images and normal-light multispectral images^[15]. In this way, the purpose of low-light image enhancement can be achieved. Compared with direct low-light image enhancement, the advantage underlying the proposed method is that the intermediate step of multispectral reconstruction will enrich the data information for low-light image enhancement. In addition, the dense connections used in the network structure not only help to alleviate the possibility of non-convergence of network training but also the number of channels in the dense blocks greatly reduces the amount of parameters. Finally, the hybrid attention mechanism of CBAM is used to weight the feature channels^[21] and reduce the patch noise in the reconstructed spectrum, so that the obtained low-light enhancement RGB image has better visual effects than that enhanced by existing enhancement methods.

1 Theory and Models

From the perspective of image decomposition, the Retinex theory is proposed to estimate the illumination components in low-light images $L(x,y)$ and reflection component $R(x,y)$. Assuming the essence of the object's color depends on the reflection component and is not affected by the illumination component, image enhancement is achieved by removing or correcting the illumination component. Assuming the image consists of an illumination component $L(x,y)$ and a reflection component $R(x,y)$, and the image $I(x,y)$ formula can be described as follows.

$$I(x, y) = R(x, y) \times L(x, y) \quad (1)$$

Three basic elements of light source, object, and camera are involved in the imaging process. Among them, a light source refers to a collection of different spectra in electromagnetic wavelengths. It is usually characterized by spectral power distribution (SPD). When the radiant energy of the light source irradiates the surface of the object, the object will selectively absorb part of the wavelength of the light source radiant energy and reflect the rest, forming a radiation spectrum that combines the light source information and the object's reflection characteristics. The radiation spectrum is focused by the camera lens and incident on the camera sensor. After photoelectric (OE) and analog-to-digital (AD) conversion, an initial raw format digital image is formed on the camera sensor, and then it goes through a series of image signal processing (ISP) steps to finally form a color image that conforms to visual perception^[13]. The entire imaging process can be divided into a linear imaging stage and a nonlinear correction stage. The initial raw format image is the linearized imaging data, and image signal processing (ISP) makes the raw image data nonlinear. For different camera manufacturers and different cameras, the algorithms used are different^[22]. Therefore, it is assumed that the ideal imaging model is represented as follows,

$$\begin{aligned}
d_i &= \int l(\lambda)r(\lambda)t(\lambda)f_i(\lambda)s(\lambda)d\lambda + n_i \\
&= \int m_i(\lambda)r(\lambda)d\lambda + n_i
\end{aligned} \tag{2}$$

where d_i is the raw response of the i -th channel of a pixel at (x,y) in the image, $l(\lambda)$ is the spectral power distribution of the lighting source, $r(\lambda)$ is the spectral reflectance of the surface of an object, $t(\lambda)$ is the overall transmittance of the camera lens optics, $f_i(\lambda)$ is the transmittance of the i -th channel filter of the camera, $s(\lambda)$ is the spectral sensitivity function of the camera sensor, λ is the wavelength, n_i is the noise signal of the i -th channel of the digital camera, $m_i = l(\lambda)t(\lambda)f_i(\lambda)s(\lambda)$ is the overall spectral sensitivity function of the i -th channel of the digital camera.

Due to the very-high correlation between RGB responses and their corresponding multispectral reflectance^[14], the mathematical algorithms can be used to directly learn the mapping relationship between RGB and multispectral images. With the success of deep learning in various computer vision tasks, convolutional neural networks have been introduced into multispectral reconstruction. Assuming that an RGB image and its corresponding multispectral image are given, the mapping between RGB and multispectral images can be described in Equation (3).

$$r = f(d) \tag{3}$$

In equation (3), it is supposed that the conversion function $f(\bullet)$ from RGB to spectral is known, that is, the feature extraction and feature selection operation of the response signal d can be realized, and the calculation of the spectral information r can be realized. Given any RGB image d , their corresponding spectral images r can be reconstructed.

2 Proposed Method

This paper proposes a low-light image enhancement method based on multispectral reconstruction. The process of this method includes two parts (see Figure 1), the construction of the low-light image enhancement module (named as modeling) and the application of the module (named as application). Three key steps are included in the modeling part, they are the multispectral reconstruction of normal-light images, the design of the deep learning architecture for low-light image enhancement, and the training of the low-light image enhancement network. In the first step of the modeling part, the normal-light RGB images in the LOw Light dataset (LOL) are used as input data^[23], and the metameric multispectral images from 400nm to 700nm with a step interval of 10nm of the normal-light RGB image are reconstructed based on reversing the ISP method that proposed by Joyce *et al.*^[24]. The multispectral reconstruction principle in the Toolbox of ISET for spectral reflectance inspection is based on the inverse imaging pipeline of digital imaging^[24], where the RGB image is inversely quantized and inversely gamma corrected to obtain the linear raw image. Then, following the inverse AD and OE conversion under the assigned illuminant (e.g., CIE D65) and camera sensitivity functions (e.g., a default in ISET), we can calculate the corresponding metameric multispectral image of the RGB image.

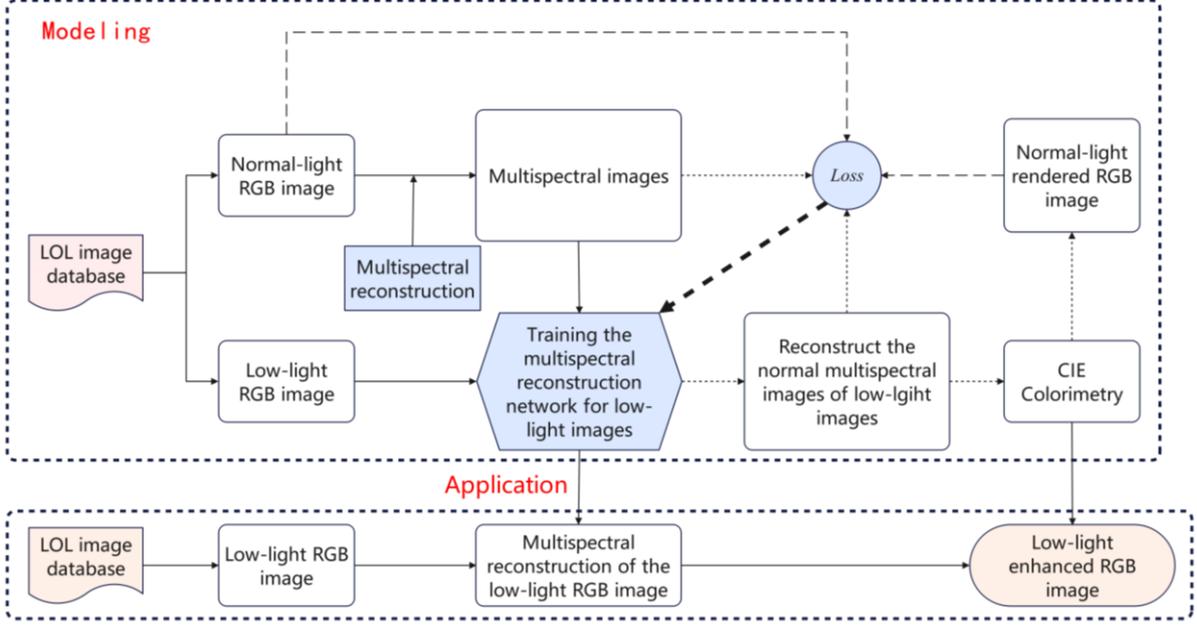


Figure 1. Flow chart of the proposed low-light image enhancement method.

In the second step of designing the deep learning architecture for low-light image enhancement, the low-light RGB images and the normal-light multispectral images reconstructed in the first step are used as training data pairs to train a deep multispectral reconstruction network. In the third step, the proposed multispectral reconstruction architecture of low-light image enhancement is trained with the paired data and the designed *Loss* functions. Finally, with the established multispectral reconstruction network, the normal-light metameric multispectral image of any low-light image can be acquired. Then the corresponding normal-light RGB image can be calculated and rendered based on the CIE colorimetry. For the calculation of the CIE XYZ tristimulus values in the proposed method, as shown in Equation (4), the CIE1964 10-degree color matching functions (CMFs) and CIE D65 illumination (after tested) are used, and the standard sRGB color space is used to represent the rendered RGB images of the original low-light images.

$$\begin{cases} X = k \int_{380}^{780} l(\lambda) r(\lambda) \bar{x}(\lambda) d\lambda \\ Y = k \int_{380}^{780} l(\lambda) r(\lambda) \bar{y}(\lambda) d\lambda \\ Z = k \int_{380}^{780} l(\lambda) r(\lambda) \bar{z}(\lambda) d\lambda \end{cases} \quad (4)$$

$$k = \frac{100}{\int_{380}^{780} r(\lambda) \bar{y}(\lambda) d\lambda}$$

where X , Y , and Z are the tristimulus values of the spectral reflectance in a pixel, $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, and $\bar{z}(\lambda)$ are standard observer color matching functions, $r(\lambda)$ is the spectral reflectance of a

pixel, $l(\lambda)$ is the light source's relative spectral power distribution function, λ is the wavelength, and k is the modulation factor. With the tristimulus in each pixel, the corresponding sRGB values are calculated using the preset 3-by-3 color conversion matrix^[25].

When designing the deep multispectral image reconstruction network (see Figure 2), some improvements are made based on the network proposed by Zhang *et al.*^[15], because the number of network parameters is less, and the hardware resources consumed during training and testing are relatively small. During the network training stage, the hyper-parameters are kept consistent with the original network. The network is trained for 50 epochs, the batch size is set to 64, and the Adam optimizer is used to train the network. The learning rate is initially set to 0.0001 and decays exponentially at a rate of 0.99. The loss function used the error between the reconstructed spectral reflectance and the groundtruth spectral reflectance, as well as the structural similarity. The composition of the loss function is shown in Equation (5),

$$Loss = \|r_{rec} - r_f\|_1 + loss_{ssim}(r_{rec}, r_f) \quad (5)$$

where r_{rec} is the reconstructed spectral reflectance, r_f is the reference spectral reflectance acquired in the previous step. $\|\cdot\|_1$ denotes the L_1 norm, the $loss_{ssim}(r_{rec}, r_f)$ is equal to one minus $ssim(r_{rec}, r_f)$, and $ssim(r_{rec}, r_f)$ represents the structural similarity (SSIM) between r_{rec} and r_f . At the same time, the optimization strategy used in our previous study^[26] is also used in this work to further improve the multispectral effect, where the convolutional block attention module (CBAM) is added to the neural network as a plug-and-play module^[21].

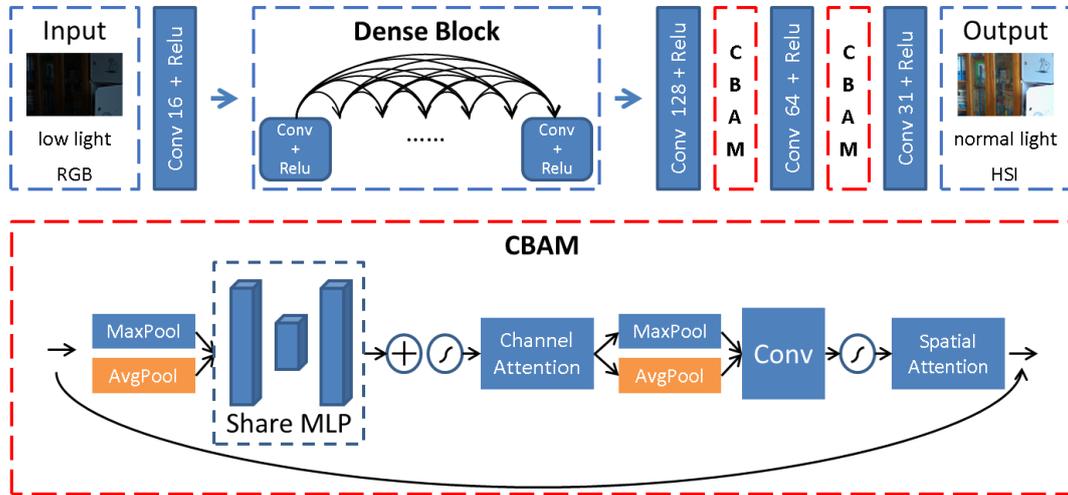


Figure 2. Architecture of the proposed optimized low-light image enhancement network.

With the designed low-light image enhancement network in Figure 2, the input low-light RGB image first passes through 16 layers of convolution to extract the shallow feature information in the image, and then passes through 7 layers of dense network modules. Each layer has 16 convolutions. Compared with the residual structure, the dense structure further improves the reuse rate of channel features. The size of the feature map remains unchanged, and the number of channels gradually increases as the network deepens, greatly reducing the parameters of the network and also mitigating the possibility of non-convergence of network

training to a certain extent. After shallow feature extraction and applying a dense connection network, a feature information map with 128 layers is finally learned, and then input into the reconstruction layer including three layers of convolution. The kernel size of each layer of convolution in the network is 3. The Relu is used as an activation function, and the CBAM module is added to the reconstruction layer to further improve the network for better image quality.

In the convolutional block attention module (CBAM), the channel attention uses average pooling and maximum pooling layers to compress feature information, and then extracts feature weights through two fully connected layers including activation. The first fully connected layer can reduce the feature dimension to $1/r$ of the input, where r is the compression parameter. In this article, r is 32 and 16 respectively for the two blocks. After the feature map output by the previous layer is activated by the *Relu* function, it is restored to the original dimension by the second fully connected layer. The feature information processed by the second fully connected layer is processed by the sigmoid function, and the original features are recalibrated in the channel domain. The spatial attention mechanism module takes the output of channel attention as input and uses average pooling and maximum pooling to integrate channel feature information, and then combines the obtained features to reduce the dimension through convolution with a convolution kernel size of 1, Then the required mask is obtained through the sigmoid activation function.

3 Experimental

The experiment was completed on a desktop computer. The processor used was Intel Core I5-13400, the graphics card used was NVIDIA GeForce RTX4060Ti, the programming environment was Python 3.8, and the TensorFlow deep learning framework was used. During the experiment, the data set used is the LOL data set.

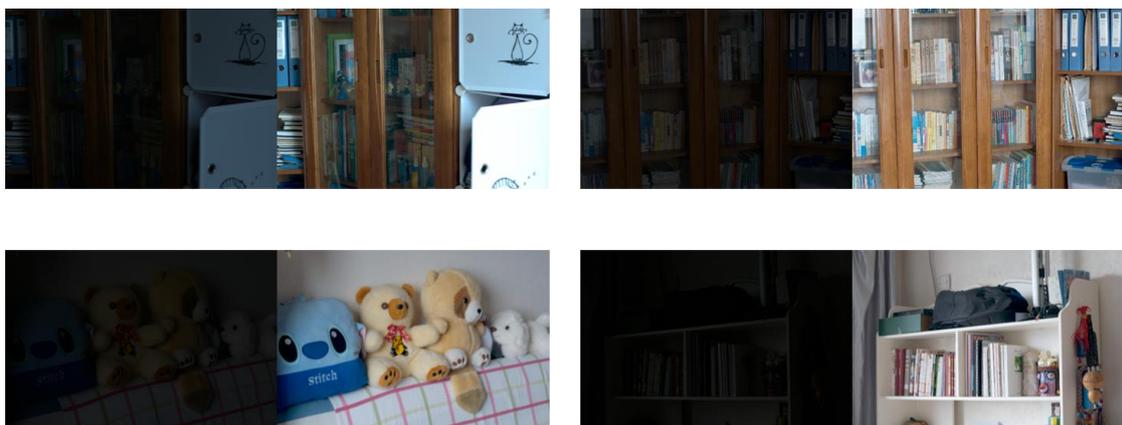


Figure 3. Some of low-light RGB images (left) and normal-light RGB images (right) in the LOL dataset.

The LOL dataset contains a data set of image pairs with a bit-depth of 8 bits of each channel obtained from real scenes for low-light enhancement tasks, which contains images of various scenes, such as houses, campuses, streets, and so on^[23]. The data set has a total of 500 pairs of data, including 485 pairs of images for training and 15 pairs of images for testing. The

resolution of each image is 400×600 pixels. The physical inverse-based multispectral reconstruction method proposed in the literature [24] is implemented in Matlab to generate the corresponding multispectral images of the normal-light images in LOL. Some of the image pairs in the LOL data set are shown in Figure 3, where the left side is the low-light images and the right is the normal-light images. During training, 485 image pairs were cropped into 40×40 pixels sample pairs and used to train the proposed network. The training process includes 50 epochs of training, with 64 pairs of samples input for each batch of training data.

4 Results and Discussion

4.1 RGB rendering under different illuminants

According to the CIE colorimetry theory and color imaging model, the multispectral image should be rendered to RGB images to make it perceptible to the human visual system. One of the important things in the proposed method is to determine which kind of illuminants should be used for RGB image rendering to reconstruct multispectral images. With the normal-light RGB images in the LOL dataset as the groundtruth, we have tested a series of illuminants during the experiment to find out the suitable illuminant for RGB image rendering. Taking one test image as an example, Figure 4 shows the rendered RGB images under different illuminants of A, D50, F7, F11, and D65.

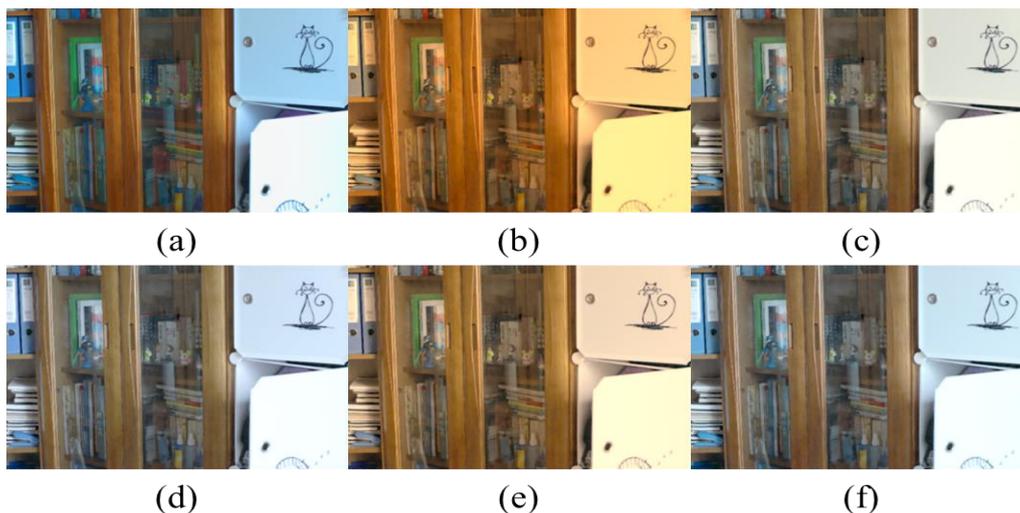


Figure 4. One of the ground-truth normal-light image in the testing sets (a), and its corresponding reconstructed images from the low-light image under different tested illuminants: (b) illuminant A, (c) illuminant D50, (d) illuminant F7, (e) illuminant F11, and (f) illuminant D65.

It can be seen from Figure 4 that the reconstructed multispectral images show different visual effects under different illuminants that with different correlated color temperatures (CCT). Generally speaking, the lower the CCT, the warmer the visual perception, while the higher the CCT, the colder the visual perception. But we can not visually determine which illuminant is the good enough from several available candidates for RGB image rendering in the proposed method. Therefore, in order to determine the most suitable illuminant from several candidates for RGB image rendering, the normal-light image was used as the reference and groundtruth,

we calculated the overall average color difference (CIEDE2000) of all 15 testing images between the rendered and the normal-light RGB images under different illuminants (see Table 1). At the same time, the metrics of the average peak signal-to-noise ratio (PSNR), structural-similarity-index-measure (SSIM) and root-mean-square error (RMSE) between the rendered and the groundtruth RGB images were also calculated and summarized in Table 1. RMSE was calculated, as shown in Equation (6), where n represents the channels of the RGB image, i represents a pixel in a image, I_R represents the rendered image, I_g represents the corresponding normal-light RGB images.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (I_R^{(i)} - I_g^{(i)})^2} \quad (6)$$

Table 1. Quantitative comparison of the RGB rendering effects of all 15 tested low-light image enhancement and the corresponding groundtruth normal-light images under different rendering illuminants.

	CIEDE2000	PSNR	SSIM	RMSE(%)
A	20.94	14.73	0.27	17.45
D50	11.09	20.89	0.62	8.74
F7	9.24	21.63	0.77	8.35
F11	14.99	18.52	0.43	11.23
D65	9.24	21.74	0.78	8.25

It is easy to find out from Table 1 that the overall average error of different metrics gives the outperform result under the illuminant D65, and then followed by illuminants F7, D50, F11, and A. This is reasonable as people have adapted the illuminant D65 more than other illuminants overall, and most of the current imaging devices perform white balance and color correction using D65 as a reference. Therefore, we chose the illuminant CIE D65 for RGB image rendering after we got the reconstructed multispectral image of a low-light RGB image in LOL dataset. It should be noted that the rendering illuminant was determined using the testing images, not the training images. As a result, the training of the multispectral reconstruction work in Figure 1 is also influenced by the testing images.

4.2 Comparison with existing methods

After determining the rendering illuminant for RGB image rendering, the performance of the proposed method is compared with some of the existing methods, such as RRM^[5], KinD^[6], Self-supervised^[8], RetinexNet^[23], RUAS^[9], URetinex-net^[10], and the original multispectral

reconstruction network of Zhang *et al.*^[15], based on the same dataset of LOL, which includes 485 training and 15 testing images. The CIEDE2000 color difference, the peak signal-to-noise ratio (PSNR), the structural similarity(SSIM), the root-mean-square error (RMSE), and the learned perceptual image patch similarity (LPIPS)^[27] are calculated to quantify the performance of each method. The average values of each metric for different methods are summarized in Table 2.

Table 2. Quantitative comparison of the low-light image enhancement performance of the proposed and some existing methods based on the average metrics error of 15 testing images.

	CIEDE2000	PSNR	SSIM	RMSE(%)	LPIPS
RetinexNet ^[23]	15.73	16.77	0.56	13.47	0.419
Self-supervised ^[8]	12.56	19.50	0.71	11.18	0.322
RRM ^[5]	20.68	13.88	0.66	21.26	0.290
KinD ^[6]	12.48	17.65	0.76	12.86	0.216
RUAS ^[9]	16.77	16.40	0.58	16.13	0.303
URetinex-net ^[10]	10.63	19.84	0.78	10.40	0.168
Zhang ^[15]	11.45	18.70	0.73	11.06	0.294
Ours	9.24	21.74	0.78	8.25	0.211

It is easy to see from Table 2 that the errors of our method for CIEDE2000, PSNR, SSIM, and RMSE are 9.24, 21.74, 0.78, and 8.25%, apparently outperforming the existing methods. For the metric of LPIPS, the best result comes from URetinex-net with a value of 0.168, and our method gives the second-best result followed by KinD, RRM, and so on. This shows that the proposed low-light image enhancement method based on multispectral reconstruction significantly outperforms the existing methods, in the case of using CIE D65 illuminant for RGB image rendering. In addition, by comparing our optimized multispectral reconstruction method with the original Zhang’s method, we can see that the optimized strategy in the proposed method can significantly improve the low-light image enhancement results, and more than 16% improvement for the evaluation metrics of CIEDE2000, PSNR, and RMSE are observed. To intuitively compare the experiment results of different methods, some low-light image enhancement results from the 15 testing images are shown in Figure 5.

We can see from Figure 5 that even different methods have improved the visual effect of the original low-light image, but when compared with the normal-light RGB images in the top-row, the RetinexNet method shows apparent image noise and color bias over the image in the enhancement results. The self-supervise method exhibits better results on noise suppression than the RetinexNet method but still has an apparent color bias on colorful images. Both the RRM and KinD methods exhibit a slight color bias, and the enhanced images still appear relatively dim compared to the normally lit RGB image in the top row.

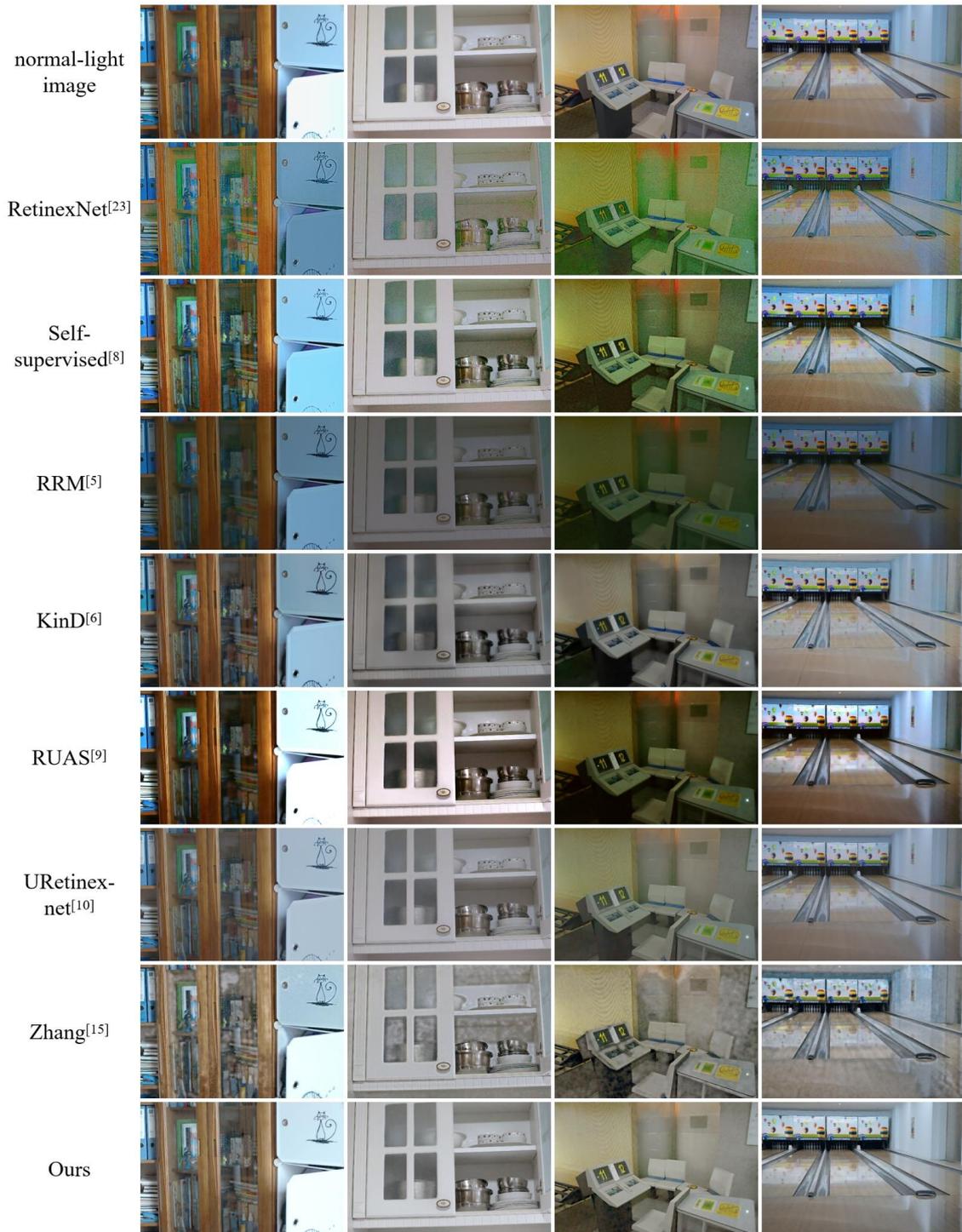


Figure 5. Some low-light image enhancement results from the 15 testing images of the proposed and the existing methods, where the first row is the norm-light images and rows 2 to 9 present enhancement low-light images.

In addition, it appears that the image saturation from RUAS is significantly enhanced, resulting in higher contrast of the images. However, this enhancement is still noticeable in relatively low light levels of the results. The URetinex-net gives the relatively good results but less saturated. The image enhanced by Zhang’s method has good brightness similar to the

normal-light image, but it contains noticeable noise patches in solid color areas. Our optimized method effectively addresses the noise patches present in Zhang's method and delivers an improved image enhancement compared to normal-light images.

Overall, compared to the normal-light RGB images in the LOL database, our method has demonstrated better results for the testing images of the LOL database than most existing methods in reducing image noise and maintaining structure and clarity. However, it still falls short in maintaining the color saturation of the images, which leads to its inferior performance compared to the URetinex-net in terms of the LPIPS metric. It is a critical issue that requires further improvement in our method.

5 Conclusion

This research is based on imaging theory and uses multispectral reconstruction to enhance low-light images. The method's effectiveness lies in the two-step process of obtaining metameric multispectral images from normal-light images and then rendering into RGB images. According to the Retinex theory, the reflection component, which equals to the reconstructed multispectral image in this study, represents the essential and unchanging color features regardless of the external environment. Thus, low-light image enhancement through multispectral reconstruction has the potential to deliver better results. Additionally, we incorporated an advanced deep learning-based multispectral reconstruction method into low-light image enhancement with CBAM to achieve better results. Experimental results have demonstrated the effectiveness of the proposed method. However, the results of test images were influenced by the use of a test image to choose the rendering illuminant, which impacted the training of the multispectral reconstruction network, and further optimization should be introduced to the proposed method to improve its performance on color quality.

Acknowledgement

This work was supported by National Natural Science Foundation of China (62305255), Hubei Provincial Natural Science Foundation General Project (No.2022CFB537), Hubei Provincial Department of Education Science and Technology Research Program Youth Talent (No.Q20221706), and China Scholarship Council (202308420128).

Conflicts of Interest

The authors declare no conflict of interest.

References

- [1] Land E H. The retinex theory of color vision. *Scientific american*, 1977, 237(6): 108-129.
- [2] Jobson D J, Rahman Z, Woodell G A. Properties and performance of a center/surround retinex *IEEE transactions on image processing*, 1997, 6(3): 451-462.
- [3] Jobson D J, Rahman Z, Woodell G A. A multiscale retinex for bridging the gap between color images and the human observation of scenes. *IEEE Transactions on Image processing*, 1997, 6(7): 965-976.
- [4] Wang S, Zheng J, Hu H M, et al. Naturalness preserved enhancement algorithm for non-uniform illumination images. *IEEE transactions on image processing*, 2013, 22(9): 3538-3548.

- [5] Li M, Liu J, Yang W, et al. Structure-revealing low-light image enhancement via robust retinex model. *IEEE Transactions on Image Processing*, 2018, 27(6): 2828-2841.
- [6] Zhang Y, Zhang J, Guo X. Kindling the darkness: A practical low-light image enhancer. *Proceedings of the 27th ACM international conference on multimedia*. 2019: 1632-1640.
- [7] Wang L W, Liu Z S, Siu W C, et al. Lightning network for low-light image enhancement. *IEEE Transactions on Image Processing*, 2020, 29: 7984-7996.
- [8] Zhang Y, Di X, Zhang B, et al. Self-supervised image enhancement network: Training with low light images only. *arXiv 2020. arXiv preprint arXiv:2002.11300*, 2020.
- [9] Liu R, Ma L, Zhang J, et al. Retinex-inspired unrolling with cooperative prior architecture search for low-light image enhancement. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2021: 10561-10570.
- [10] Wu W, Weng J, Zhang P, et al. Uretinex-net: Retinex-based deep unfolding network for low-light image enhancement. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2022: 5901-5910.
- [11] Jeon J J, Park J Y, Eom I K. Low-light image enhancement using gamma correction prior in mixed color spaces. *Pattern Recognition*, 2024, 146: 110001.
- [12] Guo X, Hu Q. Low-light image enhancement via breaking down the darkness. *International Journal of Computer Vision*, 2023, 131(1): 48-66.
- [13] Nakamura J. *Image Sensors and Signal Processing for Digital Still Cameras*. CRC Press, Inc.2005.
- [14] Lin Y T, Finlayson G D. Physically plausible spectral reconstruction from RGB images. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. 2020: 532-533.
- [15] Zhang J, Sun Y, Chen J, et al. Deep-learning-based hyperspectral recovery from a single RGB image. *Optics Letters*, 2020, 45(20): 5676-5679.
- [16] Li J, Wu C, Song R, et al. Adaptive weighted attention network with camera spectral sensitivity prior for spectral reconstruction from RGB images. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. 2020: 462-463.
- [17] Hu J, Shen L, Sun G. Squeeze-and-excitation networks. *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018: 7132-7141.
- [18] Shi Z, Chen C, Xiong Z, et al. Hscnn+: Advanced cnn-based hyperspectral recovery from rgb images. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*. 2018: 939-947.
- [19] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016: 770-778.
- [20] Huang G, Liu Z, Van Der Maaten L, et al. Densely connected convolutional networks. *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017: 4700-4708..
- [21] Woo S, Park J, Lee J Y, et al. Cbam: Convolutional block attention module. *Proceedings of the European conference on computer vision (ECCV)*. 2018: 3-19.
- [22] Ribes A, Schmitt F. Linear inverse problems in imaging. *IEEE Signal Processing Magazine*, 2008, 25(4): 84-99.
- [23] Wei C, Wang W, Yang W, et al. Deep retinex decomposition for low-light enhancement. *arXiv preprint arXiv:1808.04560*, 2018.

- [24] Farrell J, Catrysse P, Wandell B. The digital camera is an imaging system. *Imaging Systems*. Optica Publishing Group, 2010: JTUA26.
- [25] http://www.brucelindbloom.com/index.html?Eqn_RGB_XYZ_Matrix.html (access available, 07-07-2024)
- [26] Liang J, Xin L, Zuo Z, et al. Research on the deep learning-based exposure invariant spectral reconstruction method. *Frontiers in Neuroscience*, 2022, 16: 1031546.
- [27] Zhang R, Isola P, Efros AA, et al. The unreasonable effectiveness of deep features as a perceptual metric. *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018: 586-595.