https://doi.org/10.2352/lim.2024.5.1.16

Investigation on low-light image enhancement based on multispectral reconstruction

Jinxing Liang[1,4], Zhuan Zuo[1], Lei Xin[2,*], Xiangcai Ma[3], Hang Luo[1], Xinrong Hu[1], Kaida Xiao[4]; [1]Wuhan Textile University, Wuhan 430200, China; [2]Wuchang University of Technology, Wuhan 430223, China; [3]Shanghai Publishing and Printing College, Shanghai 200093, China; [4]University of Leeds, Leeds,LS2 9JT, UK

Abstract

Low-light image enhancement is a hot topic as the low-light image cannot accurately reflect the content of objects. The use of low-light image enhancement technology can effectively restore the color and texture information. Different from the traditional low-light image enhancement method that is directly from low-light to normal-light, the method of low-light image enhancement based on multispectral reconstruction is proposed. The key point of the proposed method is that the low-light image is firstly transformed to the spectral reflectance space based on a deep learning model to learn the end-to-end mapping relationship from a low-light image to a normal-light multispectral image. Then the the corresponding normal-light color image is rendered from the reconstructed multispectral image and the enhancement of the low-light image is completed. The motivation behind the proposed method is whether the low-light image enhancement through multispectral reconstruction will help to improve the enhancement performance or not. The verification of the proposed method based on the commonly used LOL dataset showed it outperforms the traditional direct enhancement methods, however, the underlying mechanism of the method is still to be further studied.

Introduction

Images captured with insufficient lighting intensity have defects such as underexposure and lack of image details. To address this problem, researchers have proposed many low-light image enhancement methods[1-8]. The Retinex model, assuming that the image can be decomposed into two components of illumination and reflectance[1], is widely used in low-light correction tasks, and some optimized methods including Single-scale and Multi-scale Retinex have been proposed[2,3]. Wang *et al.* proposed to enhance image detail while maintaining lighting naturalness[4]. Li *et al.* proposed a robust Retinex model by handling low-light image enhancement under strong noise conditions[5]. For the deep learning-based low-light enhancement, the KinD network is proposed based on decomposing images into reflection components and illumination components[6]. Li *et al.* Proposed the LightenNet to enhance low-light images[7]. In addition, Zhang *et al.* proposed a self-supervised low-light image enhancement method based on deep learning, and the network only requires minutes of training to achieve image enhancement[8].

However, most of the current methods do low-light image enhancement directly from image to image, and seldom methods do image enhancement through spectral reflectance space. As is known to all, three key factors in the theory of digital imaging model, spectral reflectance, light source, and sensitivity functions[9]. If we can reconstruct the normal-light multispectral image for the low-light image, then we can get its normal-light image under any light source and sensitivity functions. This idea can be realized when introducing the multispectral reconstruction technique[10] into the low-light image enhancement workflow. This need we can first acquire the multispectral images of the normal-light images, and then establish the end-to-end mapping models from a low-light image to a normal-light multispectral image, such as the recently reported deep learning-based multispectral reconstruction models[11-16].

Therefore, a multispectral reconstruction-based low-light image enhancement workflow is proposed in this paper (as shown in Figure 1). Based on the LOL dataset, the multispectral images of normal light RGB images are calculated using the underlying functions of Image Systems Evaluation Toolbox (ISET) for spectral data inspection[17]. Then, the dense neural network proposed by Zhang *et al.*[11] with attention mechanism[17] is developed to establish the mapping relationship between low-light RGB images and normal-light multispectral images. At last, the reconstructed multispectral images of low-light images are rendered to sRGB color space. Experimental results demonstrate the effectiveness of the proposed method.

Theory and Models

For digital imaging, the radiant energy of the light source irradiates to surface of an object, and the object will selectively absorb part of the radiant energy and reflect the rest, forming a radiation spectrum. The radiation spectrum will be focused by the camera lens and incident on the camera sensor, and after photoelectric and analog-to-digital conversion, an initial raw format digital image is formed on the camera sensor, and then goes through a series of image signal processing steps to finally form a color image that conforms to visual perception[9]. When considering only the linear stage of the digital imaging process[19], the imaging can be modeled as Eq(1):

$$d_i = \int l(\lambda) r(\lambda) s_i(\lambda) d\lambda + n_i \tag{1}$$

where d_i is the raw response of the i-th channel of a pixel, $l(\lambda)$ is the spectral power distribution of the lighting source, $r(\lambda)$ is the spectral reflectance of a point on the surface of an object, $s_i(\lambda)$ 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.Due to the high correlation between RGB values and corresponding multispectral reflectance[10], therefore the learning-based methods can be used to model multispectral images and directly learn the mapping relationship between RGB and multispectral images. Assuming that an RGB image and its corresponding multispectral image are given, the mapping between RGB and multispectral images can be described as in Equation (2).

$$R = F(D) \tag{2}$$

where F() represents the conversion model from RGB to multispectral. Given any RGB image D, their corresponding multispectral images R can be reconstructed.

Proposed Method

The process of the proposed method includes two parts, the construction of the low-light image enhancement model and the application of the model. Three key steps are included in the modeling part: 1) multispectral calculation of normal-light images, 2) the design of the deep learning architecture for low-light image enhancement, and 3) training of the low-light image enhancement model.

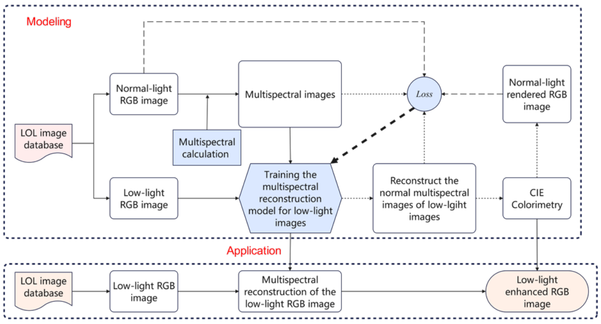

Figure 1. *Flowchart of the proposed low-light image enhancement method*

In the modeling part, the normal-light RGB images in the LOL dataset are used as input data[20], and the multispectral images from 400nm to 700nm with a step interval of 10nm are firstly calculated the underlying functions of ISET for spectral data inspection[17]. 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 are used as training data to train a deep multispectral reconstruction model. In third, the proposed multispectral reconstruction architecture of low-light image enhancement is trained. Finally the normal-light multispectral image of any low-light image can be acquired, and the corresponding sRGB image can be rendered based on CIE colorimetry using CIE1964 10-degree color matching function and CIE D65 illumination.

For the deep multispectral image reconstruction model design, the architecture established by Zhang *et al.*[11] that with less parameters is adopted with some improvements introduced[18]. During the model training stage, the hyper-parameters and the loss functions are kept consistent with the original model[11]. The composition of the loss function is shown in Equation (3),

$$\text{loss} = \left\|R_{rec} - R_{gt}\right\|_1 + \left\|I_{rec} - I_{gt}\right\|_1 + \text{loss}_{ssim}\left(I_{rec}, I_{gt}\right) \qquad (3)$$

where R_{rec} is reconstructed multispectral image, R_{gt} is groundtruth. I_{rec} and I_{gt} are rendered RGB image of reconstructed and groundtruth multispectral image. $\| \|_1$ denotes the L1 norm, the $\text{loss}_{ssim}(I_{rec},I_{gt})$ is equal to one minus $\text{ssim}(I_{rec},I_{gt})$, and $\text{ssim}(I_{rec},I_{gt})$ is the structural similarity of I_{rec} and I_{gt}. At the same time, the optimization strategy with convolutional block attention module (CBAM) we used in the literature[22] is also used in this work to improve the multispectral effect.

The architecture of the proposed multispectral reconstruction based low-light image enhancement method is shown in Figure 2.

The low-light RGB image is first processed through 16 layers of convolution to extract shallow feature information. It then goes through 7 layers of dense network modules, each containing 16 convolutions. After shallow feature extraction and dense connection networking, a feature information map with 128 layers is learned and then input into the reconstruction layer, which consists of three layers of convolution. Each convolution layer in the network has a kernel size of 3. The ReLU activation function is used, and the CBAM module is added to the reconstruction layer to further improve network robustness..

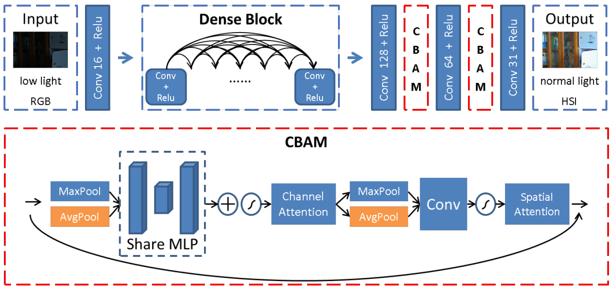

Figure 2. *Architecture of the proposed optimized low-light image enhancement model*

In the 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. 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, 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, and then through the sigmoid activation function gets the required mask.

Experimental

The experiment was conducted on a desktop computer with an Intel Core I5-13400 processor and an NVIDIA GeForce RTX4060Ti graphics card. Python 3.8 and the TensorFlow deep learning framework were used in the programming environment. The experiment utilized the LOL dataset, which contains 500 pairs of images for low-light enhancement tasks. The dataset includes images of various scenes such as houses, campuses, and streets. It consists of 485 pairs of training images and 15 pairs of testing images, each with a resolution of 400×600 pixels. The ISET functions for spectral data inspection were implemented in Matlab to generate multispectral images from the normal-light images in the LOL dataset. During the training process, the images were cropped into 40x40 pixels and the database was expanded by flipping the images. Some of the image pairs in the LOL dataset are shown in Figure 3, where the left side is the low-light images and the right is the normal-light images.

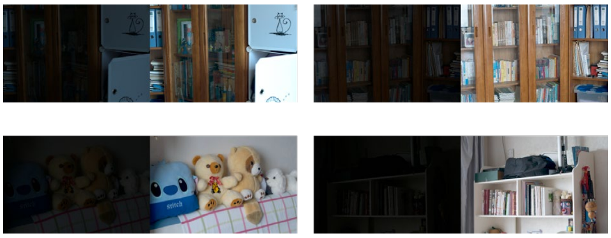

Figure 3. Some of low-light and normal-light RGB images in the LOL data set: for each pair of the image, left is the low-light image, right is the normal-light image enhancement model

Results and Discussion

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 perception 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 for the reconstructed multispectral images. With the normal-light image as the referenced groundtruth, we have tested a series of illuminants during the experiment to find out the suitable illuminant for RGB image rendering. Take one test image as an example, Figure 4 shows the RGB rendering effect under different tested illuminants of A, D50, F7, F11, and D65.

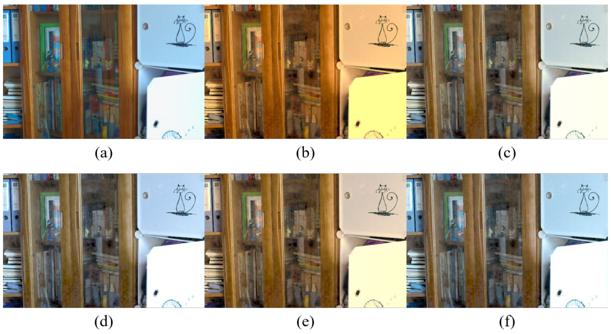

Figure 4. The RGB rendering effect under different tested illuminants of (b) A, (c) D50, (d) F7, (e) F11, and (f) D65, the first image is the normal-light image in LOL dataset

It can be seen from Figure 4 that the rendered RGB image of the reconstructed multispectral image shows different visual effects under different illuminants that with different correlated color temperatures (CCT). Generally speaking, the lower the CCT the warmer the visual perception, and the higher the CCT the colder the visual perception. However, we can not visually determine which is the best illuminant for RGB image rendering in the proposed method.

Therefore, in order to determine the most suitable illuminant for RGB image rendering, using the normal-light image as reference and groundtruth, we calculated the overall average color difference of CIEDE2000 between rendered RGB and groundtruth images under different illuminants (see Table 1). At the same time, the metrics of overall average peak signal-to-noise ratio (PSNR) and root-mean-square error (RMSE) between the rendered and the groundtruth RGB images are also calculated and summarized in Table 1.

Table 1. Quantitative comparison of the low-light image enhancement and the groundtruth normal-light images under different rendering illuminants.

	CIEDE2000	PSNR	RMSE(%)
A	20.97	14.62	17.78
D50	11.31	20.52	9.24
F7	9.71	21.19	8.81
F11	15.22	18.28	11.69
D65	9.63	21.31	8.68

It is easy to find out from Table 1 that the overall average error of different metrics gives the best 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, and most of the current imaging devices do the white balance and color correction using D65 as a white point reference. Therefore, we finnally chose the illuminant CIE D65 as for RGB image rendering after we got the reconstructed multispectral image of a low-light image.

Comparison with existing methods

After determining the rendering illuminant for RGB image, the proposed method is compared with some of the existing methods, such as RRM[5], KinD[6], Self-supervised[8], RetinexNet[20], and the original multispectral reconstruction model of Zhang *et al.*[11] to test their performance on low-light image enhancement. All the compared methods are also tested on the same dataset of LOL, which includes 485 training and 15 testing images. And also the color difference CIEDE2000, the peak signal-to-noise ratio (PSNR), and the root-mean-square error (RMSE) are used to compare the performance of each methods. Results of the low-light image enhancement of different methods are summarized in Table 2.

Table 2. Quantitative comparison of the low-light image enhancement performance of the proposed and existing methods.

	CIEDE2000	PSNR	RMSE(%)
RetinexNet	15.73	16.77	13.47
Self-supervise	12.56	19.50	11.18
RRM	20.68	13.88	21.26
KinD	12.48	17.65	12.86
Zhang	11.45	18.70	11.06
Ours	9.63	21.31	8.68

It is easy to infer from Table 2 that the overall error of our method for CIEDE2000, PSNR, and RMSE are 9.63, 21.31, and 8.68%, and apparently smaller than the existing methods. This has proven that the proposed low-light image enhancement method based on multispectral reconstruction significantly outperforms the existing methods. In addition, from comparing our optimized multispectral reconstruction method with the original Zhang's original method, we can see that the optimized strategy using CBAM in the proposed method can significantly improve the low-light image enhancement result, and more than 14% improvement for each evaluation metric is observed. To further compare the experiment results of different methods more intuitive, some low-light image enhancement results 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 images, the RetinexNet method shows apparent noise and color bias, the self-supervise method exhibits better results than the RetinexNet method but still have apparent color bias on colorful images. Both the RRM and KinD methods have slight color bias and the enhanced image still looks relatively low light. The original image enhancement result by Zhang's method has good lightness as the normal-light image, but we can see there are obvious noise patches in the solid color area. Our optimized method can solve the noise patch well in Zhang's method, and provide a more pleased image enhancement result than all of the compared method.

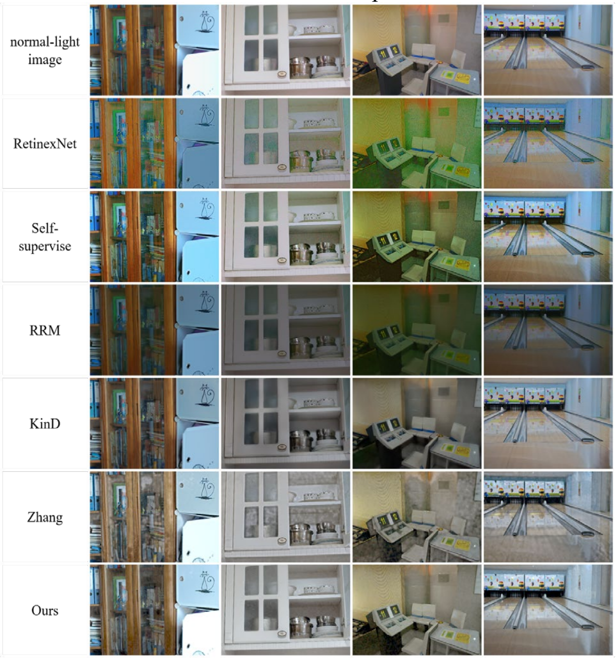

Figure 5. *Some image enhancement results of the proposed and the existing methods*

Conclusion

This research starts from imaging theory and uses spectral reflectance to achieve the goal of low-light image enhancement. The superiority of this method is mainly due to the two-step process of obtaining the multispectral images of normal-light images first and then followed by RGB image rendering. Based on the imaging theory, the reflection component that equals the reconstructed multispectral image in this study is the essence and inherent features of the color and does not change with the external environment. Therefore, low-light image enhancement via multispectral reconstruction itself has the potential to achieve better enhancement results. We also introduced the current advanced deep learning-based multispectral reconstruction method into low-light image enhancement and modified it to obtain better image enhancement. Experiments have demonstrated the superiority of the proposed method, but based on current methods, the optimization strategies for patch noise suppression in solid areas should be further studied in the future research.

Acknowledgement

This work was supported by National Natural Science Foundation of China (61575174), 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, Edwin H., The retinex theory of color vision. Sci. Am., 237, 6 (1977).
[2] Jobson D J, Rahman Z, Woodell G A., Properties and performance of a center/surround retinex, IEEE. T. Image. Process., 6,3 (1997).
[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. T. Image. Process., 6,7 (1997).
[4] Wang S, Zheng J, Hu H M, et al., Naturalness preserved enhancement algorithm for non-uniform illumination images. IEEE. T. Image. Process., 22,9 (2013).
[5] Li M, Liu J, Yang W, et al., Structure-revealing low-light image enhancement via robust retinex model. IEEE. T. Image. Process., 27,6 (2018).
[6]Zhang Y, Zhang J, Guo X., Kindling the darkness: A practical low-light image enhancer,Proc. 27th ACM ICP, pg1632. (2019).
[7] Wang L W, Liu Z S, Siu W C, et al., Lightening network for low-light image enhancement.EEE. T. Image. Process., 29,(2020).
[8] Zhang Y, Di X, Zhang B, et al., Self-supervised image enhancement network: Training with low light images only. arXiv:2002.11300, (2020).
[9] Nakamura J. Image Sensors and Signal Processing for Digital Still Cameras (CRC Press, Inc, 2005).
[10] Lin Y T, Finlayson G D., Physically plausible spectral reconstruction from RGB images. Proc. CVPR Workshops, pg.532. (2020).
[11] Zhang J, Sun Y, Chen J, et al., Deep-learning-based hyperspectral recovery from a single RGB image. Opt. Lett., 45,20 (2020).
[12]Li J, Wu C, Song R, et al., Adaptive weighted attention network with camera spectral sensitivity prior for spectral reconstruction from RGB images. Proc. CVPR Workshops, pg. 462. (2020).
[13]Hu J, Shen L, Sun G., Squeeze-and-excitation networks. Proc. CVPR, pg.7132.(2018).
[14] Shi Z, Chen C, Xiong Z, et al., Hscnn+: Advanced cnn-based hyperspectral recovery from rgb images. Proc.CVPR Workshops, pg.939.(2018).
[15] He K, Zhang X, Ren S, et al., Deep residual learning for image recognition. Proc. CVPR, Pg.770. (2016).
[16] Huang G, Liu Z, Van Der Maaten L, et al., Densely connected convolutional networks. Proc. CVPR, pg.4700. (2017).
[17] Farrell, Joyce E., Peter B. Catrysse, and Brian A. Wandell., Digital camera simulation. Appl. Optics., 51,4 (2012).
[18] Woo S, Park J, Lee J Y, et al., Cbam: Convolutional block attention module. Proc. ECCV, pg.3. (2018).
[19] Ribes A, Schmitt F., Linear inverse problems in imaging. IEEE Signal. Proc. Mag., 2008, 25,4 (2008).
[20] Wei C, Wang W, Yang W, et al., Deep retinex decomposition for low-light enhancement. arXiv:1808.04560, (2018).
[21] Liang J, Xin L, Zuo Z, et al., Research on the deep learning-based exposure invariant spectral reconstruction method. Front. Neurosci., 16, (2022).

Author Biography

Jinxing Liang is currently an associate professor in School of Computer Science and Artificial Intelligence as Wuhan Textile University, he was a doctoral students jointly trained by Wuhan University and the University of Leeds, and received his PhD in Color and Imaging Technology from Wuhan University. Now, he is a academic visitor at School of Design of University of Leeds. His research interests include multispectral reconstruction, color science, image process, knitted clothing simulation, and data-driven design.