LOW-LIGHT IMAGE ENHANCEMENT USING CNN AND BRIGHT CHANNEL PRIOR

Li Tao[1], Chuang Zhu[1*#], Jiawen Song[1], Tao Lu[2], Huizhu Jia[1#], Xiaodong Xie[1]*

[1] National Engineering Laboratory for Video Technology
Peking University, Beijing, China

[2] School of Electronic and Computer Engineering
Shenzhen Graduate School
Peking University, Shenzhen, China

ABSTRACT

In this paper, we propose a joint framework to enhance images under low-light conditions. First, a convolutional neural network (CNN) based architecture is proposed to denoise low-light images. Then, based on atmosphere scattering model, we introduce a low-light model to enhance image contrast. In our low-light model, we propose a simple but effective image prior, bright channel prior, to estimate the transmission parameter; besides, an effective filter is designed to adaptively estimate environment light in different image areas. Experimental results demonstrate that our method achieves superior performance over other methods.

Index Terms— low-light image enhancement, CNN, low-light model, bright channel prior

1. INTRODUCTION

With the popularity of smart phones and digital cameras, images play an important role in daily life. However, when images are captured in low-light conditions, usually at night, the image quality descends rapidly because of noise and loss of contrast. For tasks such as object detection, image retrieval, etc., low-light images perform badly [1]. Thus, low-light image enhancement is very important. Various enhancement methods were proposed and they can be classified into three categories: methods based on retinex theory, histogram equalization (HE) algorithms and methods using dehaze model.

Retinex theory [2] is introduced by Land and McCann. It assumes that an image is an interaction of illumination and reflectance. By removing the effect of illumination, low-light images can be enhanced. Based on this theory, several methods were proposed. Single-scale retinex (SSR) [3], multiscale retinex (MSR) [4] and multiscale retinex with color restoration (MSRCR) [5] enhance images in frequency domain while McCann99 retinex and Frankle-McCann retinex [6] process images in spatial domain. Since retinex-based methods process images in RGB channels separately, they may cause color distortion when the original image does not follow with the gray world assumption [7].

HE algorithms mainly focus on enhancing image contrast. Low-light images are usually of low-dynamic-range. By rearranging pixel values, recovered images will have stronger contrast and more proper luminance than the original ones. Traditional HE algorithm is useful for many low-dynamic-range images. However, this method tends to produce undesirable artifacts [8]. Methods like dynamic histogram equalization (DHE) [8] and brightness preserving dynamic histogram equalization (BPDHE) [9] are proposed to address this problem, however, the details in dark areas are not enhanced appropriately [10]. Still, they cause color distortion in many cases.

Recently, methods based on dehaze model achieve state-of-the-art enhancement performance. Dong et.al [11] found that after inverting, low-light images look like hazy pictures. Applying dehaze method on these inverted low-light images can achieve good performance. Similarly, Zhang et.al [12] and Li et.al [13] utilized dark channel prior [14] to do low-light image enhancement. However, these methods are lack of an established theory. Moreover, the saturation of the processed images is usually exaggerated and sometimes over-enhanced, which make results look unreal.

To solve these problems mentioned above, we formulate a low-light model based on the atmosphere scattering model, in which the light conditions at night are taken into consideration. We propose a simple but effective image prior, bright channel prior, to estimate parameters in our low-light model. Besides, a CNN based model is designed to denoise low-light images to improve image quality before the model-based enhancement.

2. IMAGE ENHANCEMENT FRAMEWORK

Inspired by super-resolution convolutional networks and atmosphere scattering model, we propose a joint approach to enhance low-light images. The first stage is using an image-in-image-out network to remove image noise. In the second stage, we apply our low-light model on the noise-free images. The framework of our method is shown in Fig. 1.

This work was partially supported by grant from the National Science Foundation of China (61602011 and 61421062), the China Postdoctoral Science Foundation 2016M590020, China 863 project of No.2015AA015903.

* These authors contributed equally to this work.
# Corresponding author (Email: {czhu, hzjia}@pku.edu.cn).

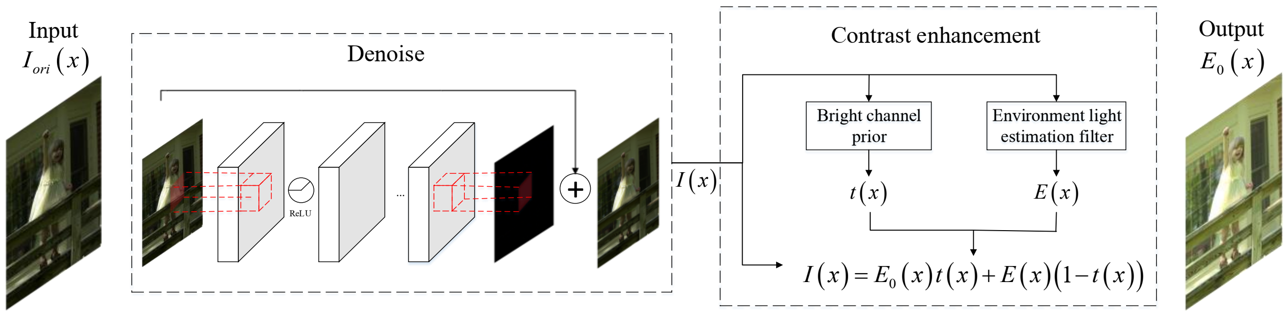

Fig. 1. Framework of our method. Given a noisy low-light image, we first use a denoising convolutional neural network to remove image noise and set the noise-free image as the input of the next stage. Then we use our low-light model to enhance image contrast.

3. THE PROPOSED METHOD

3.1. CNN model for denoising

Most commonly, noise signal in images is assumed to be independent, identically distributed [15]. Thus, denoise methods [16] [17] usually focus attention on the problem of attenuating additive white Gaussian noise (AWGN). These methods assume that the standard deviation σ of the AWGN has been accurately estimated. Many methods [18] [19] do estimate the variance accurately in most cases but still have non-negligible errors in some special situations.

Inspired by a super-resolution convolutional network (VDSR [20]) and a denoising network (DnCNN [21]), we propose a denoising convolutional neural network (DCNN) to remove image noise. By using DCNN, we don't need to calculate the standard deviation of AWGN. Meanwhile, it can get comparable results with BM3D.

3.1.1. Model

Since pictures captured by cameras are usually very large, to optimize our training process, we cut images into patches with size 41×41.

The architecture of DCNN is illustrated on the left side of Fig. 1. Our model uses 20 convolutional layers. For each layer, kernel size is 3×3 and we pad every patch in each layers during training so that the output is of the same size as the input. After convolutional layers, output is a residual image and we add it with the noise image to get a noise-free image.

3.1.2. Training

For DCNN, input and output are both images. Noise-free images are set to be ground truth and images with AWGN are input data. We utilize PCA noise level estimator [18] to estimate the standard deviation in low-light images. According to our experiment, most variances of AWGN are in range [0, 0.1], where pixel values are normalized to be in range [0, 1].Therefore, when preparing training data, we restrain the variance to be in the range [0, 0.1].

Our model is trained using Caffe [22], we set base learning rate 0.01, momentum 0.9, and weight decay 0.0001.

3.1.3. Denoising performance

The results are shown in Fig. 2. Our method get comparable results with BM3D. In BM3D, the sigma is estimated by PCA noise level estimator. By using convolutional neural network, noise is removed apparently. When we apply DCNN on low-light images (Fig. 3), the effect is still very good.

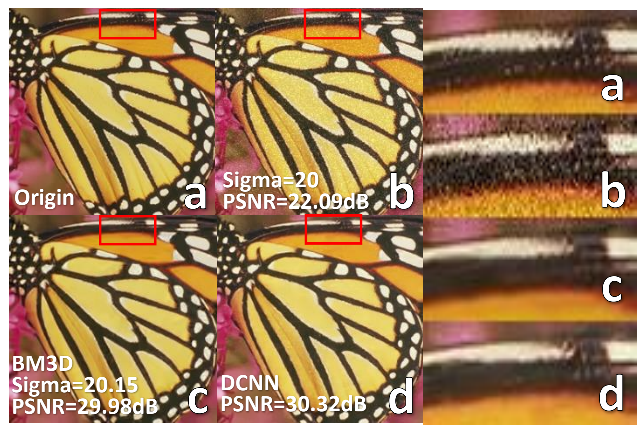

Fig. 2. The performance comparison between DCNN and BM3D.

Fig. 3. The performance of DCNN for low-light image. The image on the right is generate by the noisy one on the left using DCNN.

3.2. Contrast enhancement using proposed model

In this section, a low-light model based on atmosphere scattering model [23] is introduced. We propose bright channel prior and then use this prior in our low-light model. A filter is used to estimate environment light.

3.2.1. Low-light model

Atmosphere scattering model is widely used to describe the formation of a hazy image. The equation is as follows:

$$E(d,\lambda)=E_0(\lambda)e^{-\beta(\lambda)d}+E_\infty(\lambda)\left(1-e^{-\beta(\lambda)d}\right) \tag{1}$$

where $E(d,\lambda)$ is the image captured by camera, $E_0(\lambda)$ is the intensity of original objects without any attenuation, $\beta(\lambda)$ is the total scattering coefficient, d is the distance between the observer and the object and $E_\infty(\lambda)$ is the airlight radiance.

Based on (1), we propose a low-light model with the following equation:

$$E(d,\lambda)=E_0(\lambda)e^{-\beta(\lambda)d'}+E(x)\left(1-e^{-\beta(\lambda)d'}\right) \tag{2}$$

where $d'=d_e+d$. Then equation (2) can also be written as:

$$\begin{aligned}E(d,\lambda)&=E_0(\lambda)e^{-\beta(\lambda)d_e}e^{-\beta(\lambda)d}\\&+E(x)\left(1-e^{-\beta(\lambda)d_e}e^{-\beta(\lambda)d}\right)\end{aligned} \tag{3}$$

Assume that there is no attenuation, the true scene light is $E_0(\lambda)e^{-\beta(\lambda)d_e}$, d_e is the equivalent distance to simulate the attenuation of light intensity from daylight to low-light. Then, our target becomes recovering $E_0(\lambda)$ from $E(d,\lambda)$, instead of enhancing a low-light image by just removing the attenuation and airlight in low-light scene.

We rewrite the equation (3) as

$$I(x)=E_0(x)t(x)+E(x)\left(1-t(x)\right) \tag{4}$$

where

$$\begin{aligned}I(x)&=E(d,\lambda)\\t(x)&=e^{-\beta(\lambda)d_e}e^{-\beta(\lambda)d}\\E_0(x)&=E_0(\lambda)\end{aligned} \tag{5}$$

Though dark channel prior (DCP) [14] works on natural haze-free images, for low-light images, pixel values are all very low. Therefore, we can't use DCP to enhance low-light images in our model. We introduce a new prior to estimate $t(x)$ in Section 3.2.2.

In dehaze domain, the atmosphere light is global constant and is usually calculated using the brightest pixels [13] [14]. This doesn't work in low-light conditions, because in low-light scenes, light cannot be described as parallel light as in daylight and it is not as strong as sunlight. At night, several point light sources may exist. Shadows are different as well. Therefore, we use $E(x)$ to denote the environmental light and estimate it using an area adaptive method. The method to calculate $E(x)$ will be detailed in Section 3.2.3.

3.2.2. Estimation of $t(x)$ based on bright channel prior

Fig. 4. Bright channels of natural images. For images in sunlight, bright channels are close to 1 when pixel values are normalized in range [0, 1]. Shadow areas will have lower bright channel values.

For a single pixel, we define its bright channel the highest value in a surrounded patch $\Omega(x)$ within all color channels.

$$E_0^{bright}(x)=\max_{y\in\Omega(x)}\left(\max_{c\in\{r,g,b\}}E_0^{\,c}(y)\right) \tag{6}$$

Bright channel prior is based on the observation of outdoor images taken in sunlight conditions without any shadows. We find that for natural outdoor images, in image patches where there is no shadow, the highest value among three channels is close to 1 (the pixel values has been normalized to the range [0, 1]). If there are shadows in the image, the bright channel of that area gets darker (Fig. 4).

Thus, sunlight images have the following regularizations

$$\max_{y\in\Omega(x)}\left(\max_{c\in\{r,g,b\}}E_0^{\,c}(y)\right)\rightarrow 1 \tag{7}$$

or

$$\min_{y\in\Omega(x)}\left(\min_{c\in\{r,g,b\}}\left(1-E_0^{\,c}(y)\right)\right)\rightarrow 0 \tag{8}$$

Equation (8) suggests that the dark channel of an inverted image is close to zero, this also indicates the effectiveness of those dehaze-model-based methods [11] [12] [13].

In a local patch, $t(x)$ and $E(x)$ change a little. Thus, we assume the transmission and the environment light is constant in a small patch, and denote them as $\tilde{t}(x)$ and $\tilde{E}(x)$. We calculate the bright channel on both sides of our model:

$$\begin{aligned}\max_{y\in\Omega(x)}\left(\max_{c\in\{r,g,b\}}I(y)\right)&=\max_{y\in\Omega(x)}\left(\max_{c\in\{r,g,b\}}E_0(y)\right)\tilde{t}(x)\\&+\left(1-\tilde{t}(x)\right)\tilde{E}(x)\end{aligned} \tag{9}$$

Since we have equation (7), transmission can be estimated by

$$\tilde{t}(x)=\frac{\max_{y\in\Omega(x)}\left(\max_{c\in\{r,g,b\}}I(y)\right)-\tilde{E}(x)}{1-\tilde{E}(x)} \tag{10}$$

This can also be written in the following form

$$\tilde{t}(x)=1-\frac{1-\max_{y\in\Omega(x)}\left(\max_{c\in\{r,g,b\}}I(y)\right)}{1-\tilde{E}(x)} \tag{11}$$

If we enhance low-light areas thoroughly, the image may seem unnatural. Therefore, we introduce a constant parameter to modify the transmission in our low-light model.

$$\tilde{t}(x)=1-\omega\cdot\frac{1-\max_{y\in\Omega(x)}\left(\max_{c\in\{r,g,b\}}I(y)\right)}{1-\tilde{E}(x)} \tag{12}$$

Here, the parameter ω can keep a small amount of the scene effect. In our method, we set ω to a fixed value of 0.8.

3.2.3. Adaptive environment light estimation

Low-light images are usually captured under non-uniform light conditions. Environment light should be adjusted according to different areas. Here, we utilize a local patch to estimate the environment light of the central pixel. We apply the following 7×7 block filter to do so.

1	1	1	1	1	1	1
1	4	4	4	4	4	1
1	4	16	16	16	4	1
1	4	16	64	16	4	1
1	4	16	16	16	4	1
1	4	4	4	4	4	1
1	1	1	1	1	1	1

Compared to average filtering and median filtering, this filter performs better in texture areas. After filtering and normalization, we get the environment light $E(x)$.

We find that $E(x)$ can adaptively enhance the image but it is a little weak, so we modify $E(x)$ as

$$E'(x) = E(x)/2 \tag{13}$$

This is easy to understand because in (13), $E'(x)$ is much closer to 0 than $E(x)$, which means in our low-light model, the environment light is closer to zero. The darker area will be enhanced more.

4. EXPERIMENTAL RESULTS

In this section, results are produced by current popular enhancement methods, SSR [3] and MSR [4] represent retinex theory, DHE [8] stands for histogram equalization and Li's method [13] is using dehaze model, along with our proposed method based on our low-light model.

All methods are tested on 42 images, which are caught in different light conditions. We set sigma 15 in SSR and 15, 80, 250 in MSR.

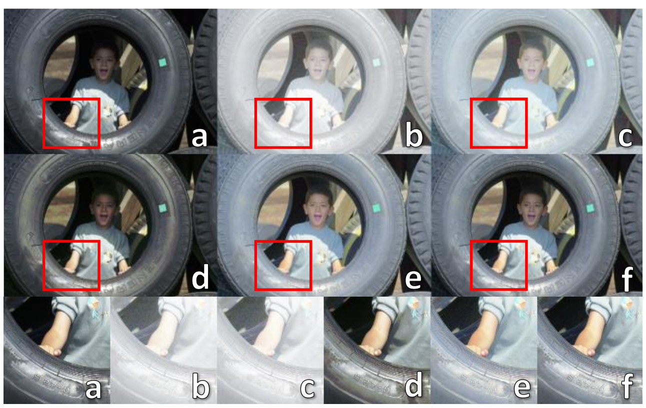

Fig. 5. Results are shown in orders. (a) is the original image, (b) to (f) respectively denote the results using different methods as SSR, MSR, DHE, Li's and the proposed.

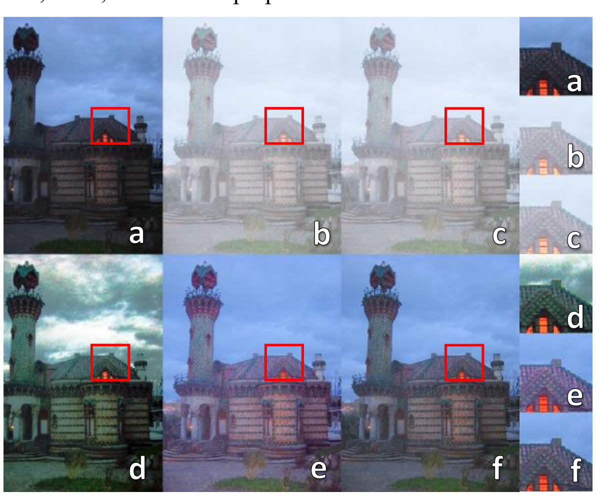

Fig. 6. Results are shown in orders. (a) is the original image, (b) to (f) respectively denote the results using different methods as SSR, MSR, DHE, Li's and the proposed.

For enhancement issues, there is usually no ground truth, therefore, it remains a problem how to properly evaluate the performance. We mainly use subjective evaluation to assess results, along with objective evaluation.

4.1. Subjective evaluation

As can be seen in Fig. 5, SSR and MSR cause color distortion while DHE even makes the image a little darker. Though Li's method doesn't have these problems, in its results, images are over-enhanced and image contrast is not recovered well. In areas with textures, details are hard to identify. In Fig. 6, DHE over enhances image contrast and changes image color style while Li's method is a little weaker in contrast enhancement. Our approach makes a good balance between contrast enhancement and texture preserving.

4.2. Objective evaluation

Here we adopt lightness-order-error (LOE) [10] and statistical naturalness measure (SNM) [24] to assess enhanced results. The smaller a LOE value is, the better the lightness order is preserved, which also means that the better the enhancement method is. As for SNM, higher results indicate that the enhanced image appears more natural.

We also use SSIM [25] to evaluate the similarity in structure. More structures are preserved when the SSIM value is higher. Results are shown in Table 1.

Table 1. Performance of different methods on 42 images.

	SSR	MSR	DHE	Li's	Proposed
LOE	46.27	45.63	25.97	34.24	**23.34**
SNM	0.157	0.156	0.413	0.489	**0.497**
SSIM	0.647	0.648	**0.933**	0.751	0.846

Though DHE don't have good visual performance, its SSIM is high because it almost doesn't change relative relationship among pixels values. Even so, our proposal method still gets the smallest LOE value and the highest SNM value among these methods. As to SSIM, we compare our results with those by DHE and find that for most images, DHE results have serious color distortion problem.

5. CONCLUSION

In this paper, a joint effective method is proposed by combining denoising and contrast enhancement for low-light images. We design a convolutional neural network to remove image noise which achieves a good performance. A low-light model is proposed and bright channel prior is used to calculate transmission in our low-light model. Furthermore, we propose an effective method to estimate environment light adaptively. Experimental results show that our algorithm can better improve brightness, enhance contrast and preserve details compared with other methods.

6. REFERENCE

[1] Huang, Kaiqi, et al. “A real-time object detecting and tracking system for outdoor night surveillance,” *Pattern Recognition*, vol. 41, pp. 432-444, 2008.

[2] Land, Edwin H., and J. J. McCann. “Lightness and Retinex Theory,” *Journal of the Optical Society of America*, vol. 61, pp. 1-11, 1971.

[3] D. J. Jobson, Z. Rahman, and G. A. Woodell, “Properties and performance of a center/surround retinex,” *IEEE Transactions on Image Processing*, vol. 6, no. 3, pp. 451–462, Mar. 1996.

[4] Z. Rahman, D. J. Jobson, and G. A. Woodell, “Multi-scale retinex for color image enhancement,” in *Proc. Int. Conf. Image Process*., pp. 1003–1006, Sep. 1996.

[5] Jobson, D. J., Z. 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 A Publication of the IEEE Signal Processing Society*, vol. 6, no. 7, pp. 965-976, 1997.

[6] Funt, Brian V, Ciurea, Florian, and McCann, John J. “Retinex in Matlab,” *Color and Imaging Conference*, pp. 112-121, 2000.

[7] Wang, Wen, et al. “A fast Multi-Scale Retinex algorithm for color image enhancement,” *International Conference on Wavelet Analysis and Pattern Recognition IEEE*, pp. 80-85, 2008.

[8] Abdullah-Al-Wadud, M, et al. “A Dynamic Histogram Equalization for Image Contrast Enhancement,” *IEEE Transactions on Consumer Electronics*, vol. 53, pp. 593-600, 2007.

[9] H. Ibrahim and N. Kong, “Brightness preserving dynamic histogram equalization for image contrast enhancement,” *IEEE Trans. Consumer Electronics*, vol. 53, no. 4, pp. 1752–1758, Nov. 2007.

[10] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: From error measurement to structural similarity,” *IEEE Transactions on Image Processing*, vol. 13, no. 1, Jan. 2004.

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

[12] Zhang, Xiangdong, et al. “Enhancement and noise reduction of very low light level images,” *Proc. International Conference on Pattern Recognition,* pp. 2034-2037, 2012.

[13] Li, Lin, et al. “A low-light image enhancement method for both denoising and contrast enlarging,” *IEEE International Conference on Image Processing IEEE*, pp. 3730-3734, 2015.

[14] K. He, J. Sun, and X. Tang, “Single image haze removal using dark channel prior,” *Pattern Analysis and Machine Intelligence (PAMI), IEEE Transactions on*, vol. 33, no. 12, pp. 2341–2353, 2011.

[15] Tian, Jing, and L. Chen. “Image Noise Estimation Using A Variation-Adaptive Evolutionary Approach,” *IEEE Signal Processing Letters*, vol. 19.7, pp. 395-398, 2012.

[16] K. Dabov, A Foi, V. Katkovnik, and K. Egiazarian, “Image denoising with block-matching and 3d filtering,” *Electronic Imaging 2006, International Society for Optics and Photonics*, pp. 606414–606414, Feb. 2006.

[17] Tomasi, C, and R. Manduchi. “Bilateral Filtering for Gray and Color Images,” *International Conference on Computer Vision IEEE*, pp. 839-846, 1998.

[18] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, “Image denoising by sparse 3-d transform-domain collaborative filtering,” *Image Processing, IEEE Transactions on*, vol. 16, no. 8, pp. 2080–2095, 2007.

[19] S. Pyatykh, J. Hesser, and L. Zheng, “Image Noise Level Estimation by Principal Component Analysis,” *IEEE Transactions on Image Processing*, vol. 22, issue: 2, Pages: 687 - 699, February 2013.

[20] Jiwon Kim, Jung Kwon Lee and Kyoung Mu Lee, “Accurate Image Super-Resolution Using Very Deep Convolutional Networks,” *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1646-1654, 2016

[21] Zhang, Kai, et al. “Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising,” *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society,* pp. 99:1-1, 2017.

[22] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell. “Caffe: Convolutional architecture for fast feature embedding,” In *Proc. of the ACM International Conf. on Multimedia*, 2014.

[23] Nayar, Shree K., and S. G. Narasimhan. “Vision in Bad Weather,” *The Proceedings of the Seventh IEEE International Conference on Computer Vision DBLP*, vol. 2, pp. 820-827, 1999.

[24] Yeganeh, Hojatollah, and Zhou Wang. “Objective quality assessment of tone-mapped images.” *IEEE Transactions on Image Processing*, vol. 22(2): 657-667, 2013.

[25] S. Wang, J. Zheng, H. M. Hu, and B. Li, “Naturalness preserved enhancement algorithm for non-uniform illumination images,” *Image Processing, IEEE Transactions* on, vol. 22, no. 9, pp. 3538–3578, 2013.