A New Image Contrast Enhancement Algorithm Using Exposure Fusion Framework

Zhenqiang Ying[1], Ge Li[1(✉)], Yurui Ren[1,2], Ronggang Wang[1], and Wenmin Wang[1]

[1] School of Electronic and Computer Engineering, Shenzhen Graduate School, Peking University, Shenzhen, China
geli@ece.pku.edu.cn

[2] Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian, China

Abstract. Low-light images are not conducive to human observation and computer vision algorithms due to their low visibility. Although many image enhancement techniques have been proposed to solve this problem, existing methods inevitably introduce contrast under- and over-enhancement. In this paper, we propose an image contrast enhancement algorithm to provide an accurate contrast enhancement. Specifically, we first design the weight matrix for image fusion using illumination estimation techniques. Then we introduce our camera response model to synthesize multi-exposure images. Next, we find the best exposure ratio so that the synthetic image is well-exposed in the regions where the original image under-exposed. Finally, the input image and the synthetic image are fused according to the weight matrix to obtain the enhancement result. Experiments show that our method can obtain results with less contrast and lightness distortion compared to that of several state-of-the-art methods.

synthetic image- expose the underexposed area

weight matric -by illumination estimation technique

input image + synthetic image arefused based on weight matrix

Keywords: Image enhancement · Contrast enhancement · Exposure compensation · Exposure fusion

1 Introduction

Image enhancement techniques are widely used in image processing. In general, it can make the input images look better and be more suitable for specific algorithms [6,14]. Contrast enhancement, as one kind of enhancement techniques, can reveal the information of the under-exposed regions in an image. Many contrast enhancement techniques have been proposed including histogram-based [3,6,12,13], Retinex-based [7,9] and dehaze-based methods [4,5].

A color image can be represented as a three-dimensional array. The simplest scheme of contrast enhancement performs the same processing for each element. For example, the earliest image enhancement methods use a non-linear monotonic function (power-law [2], logarithm [11], etc.) for gray-level mapping.

M. Felsberg et al. (Eds.): CAIP 2017, Part II, LNCS 10425, pp. 36–46, 2017.
DOI: 10.1007/978-3-319-64698-5_4

Fig. 1. Example of images that only differ in exposure.

Considering the uneven distribution of elements in different gray levels, histogram equalization (HE) is widely used for improving contrast. Many extensions to HE are proposed to take some restrictions into account such as brightness preservation [3,6,13] and contrast limitation [12]. However, HE-based methods always suffer from over-enhancement and lead to unrealistic results. Imitating the human visual system, Retinex theory is also widely used in image enhancement. By separating reflectance from the illumination, Retinex-based algorithms can enhance the details obviously [5,7,9]. However, these methods suffer from halo-like artifact in the high contrast region. In recent years, some methods borrowed the de-haze technique to contrast enhancement, and achieve good subjective visual effect [4]. But these work may cause color distortion because of contrast over-enhancement.

Although image contrast enhancement methods have been studied for decades, the definition of a good enhancement result is still not well-defined. In addition, no reference is provide for existing low-light enhancement algorithm to locate the over- and under-enhancement region. We noticed that the images that only differ in exposures can be used as a reference for enhancement algorithms, as shown in Fig. 1. As the exposure increases, some under-exposed regions become well-exposed. The enhancement result should keep the well-exposed regions unchanged and enhance the under-exposed region. Meanwhile, the contrast of enhanced regions should be consistent with the reference image that correct exposes the region.

In this paper, we proposed a new framework to help mitigate under- and over-enhancement problem. Our framework is based on exposure fusion among multi-exposure images synthesized from the input image by the camera response model. Based on our framework, we proposed an enhancement algorithm that can obtain results with less contrast and lightness distortion compared to several state-of-the-art methods.

2 Our Approach

2.1 Exposure Fusion Framework

In many outdoor scenes, cameras can not make all pixel well-exposed since its dynamic range is limited. As shown in the Fig. 1, although we can reveal

some under-exposed regions by increasing exposure, the well-exposed regions may become over-exposed at the same time. To obtain an image with all pixel well-exposed, we can fuse these images:

$$\mathbf{R}^c = \sum_{i=1}^{N} \mathbf{W}_i \circ \mathbf{P}_i^c, \tag{1}$$

where N is the number of images, $\mathbf{P}_i$ is the i-th image in the exposure set, $\mathbf{W}_i$ is the weight map of the i-th image, c is the index of three color channels and $\mathbf{R}$ is the enhanced result. It is equal for three color components and nonuniform for all pixels: the well-exposed pixels are given a big weight while the poor-exposed pixels are given a small weight. The weight is normalized so that $\sum_{i=1}^{N} \mathbf{W}_i = \mathbf{1}$.

The problem is that images with another exposure setting is not available for image enhancement problem. Fortunately, pictures taken with different exposures are highly correlated. In our earlier work, we propose a camera response model to accurately describe the association between these images so that we can generate a series of images from the input image. The mapping function between two images that only differ in exposure is called Brightness Transform Function (BTF). Given exposure ratio k_i and BTF g, we can map the input image $\mathbf{P}$ to the i-th image in the exposure set as

$$\mathbf{P}_i = g(\mathbf{P}, k_i). \tag{2}$$

In this paper, we only fuse the input image itself with another exposure to reduce complexity, as shown in Fig. 2. The fused image is defined as

$$\mathbf{R}^c = \mathbf{W} \circ \mathbf{P}^c + (\mathbf{1} - \mathbf{W}) \circ g(\mathbf{P}^c, k) \tag{3}$$

The enhancement problem can be divided into three parts: the estimation of $\mathbf{W}$, g and k. In the following subsections, we solve them one by one.

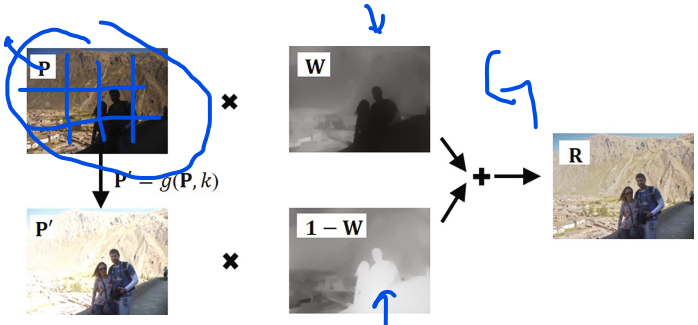

Fig. 2. Our framework.

2.2 Weight Matrix Estimation

The design of **W** is key to obtaining an enhancement algorithm that can enhance the low contrast of under-exposed regions while the contrast in well-exposed regions preserved. We need to assign big weight values to well-exposed pixels and small weight values to under-exposed pixels. Intuitively, the weight matrix is positively correlated with the scene illumination. Since highly illuminated regions have big possibility of being well-exposed, they should be assign with big weight values to preserve their contrast. We calculate the weight matrix as

$$\mathbf{W} = \mathbf{T}^{\mu} \tag{4}$$

where **T** is the scene illumination map and μ is a parameter controlling the enhance degree (See more details in Sect. 3.1). The estimation of scene illumination map **T** is solved by optimization.

Optimization Problem. The lightness component can be used as an estimation of scene illumination. We adopt the lightness component as the initial estimation of illumination:

$$\mathbf{L}(x) = \max_{c \in \{R,G,B\}} \mathbf{P}_c(x) \tag{5}$$

for each individual pixel x. Ideal illumination should has local consistency for the regions with similar structures. In other words, **T** should keep the meaningful structures of the image and remove the textural edges. As in [5], we refine **T** by solving the following optimization equation:

$$\min_{\mathbf{T}} ||\mathbf{T} - \mathbf{L}||_2^2 + \lambda ||\mathbf{M} \circ \nabla \mathbf{T}||_1, \tag{6}$$

where $||*||_2$ and $||*||_1$ are the ℓ_2 and ℓ_1 norm, respectively. The first order derivative filter ∇ contains $\nabla_h \mathbf{T}$ (horizontal) and $\nabla_v \mathbf{T}$ (vertical). **M** is the weight matrix and λ is the coefficient. The first term of this equation is to minimize the difference between the initial map **L** and the refined map **T**, while the second term maintains the smoothness of **T**.

The design of **M** is important for the illumination map refinement. A major edge in a local window contributes more similar-direction gradients than textures with complex patterns [15]. Therefore, the weight in a window that contains meaningful edges should be smaller than that in a window only containing textures. As a result, we design the weight matrix as

$$\mathbf{M}_d(x) = \frac{1}{|\sum_{y \in \omega(x)} \nabla_d \mathbf{L}(y)| + \epsilon}, \quad d \in \{h, v\}, \tag{7}$$

where $|*|$ is the absolute value operator, $\omega(x)$ is the local window centered at the pixel x and ϵ is a very small constant to avoid the zero denominator.

Closed-Form Solution. To reduce the complexity, we approximate Eq. 6 as in [5]:

$$\min_{\mathbf{T}} \sum_{x} \left(\big(\mathbf{T}(x) - \mathbf{L}(x)\big)^2 + \lambda \sum_{d \in \{h,v\}} \frac{\mathbf{M}_d(x)\big(\nabla_d \mathbf{T}(x)\big)^2}{|\nabla_d \mathbf{L}(x)| + \epsilon} \right). \tag{8}$$

As can be seen, the problem now only involves quadratic terms. Let $\mathbf{m}_d$, $\mathbf{l}$, $\mathbf{t}$ and $\nabla_d \mathbf{l}$ denote the vectorized version of $\mathbf{M}_d$, $\mathbf{L}$, $\mathbf{T}$ and $\nabla_d \mathbf{L}$ respectively. Then the solution can be directly obtained by solving the following linear function.

$$(\mathbf{I} + \lambda \sum_{d \in \{h,v\}} (\mathbf{D}_\mathbf{d}^\intercal Diag(\mathbf{m}_d \oslash (|\nabla_d \mathbf{l}| + \epsilon))\mathbf{D_d})\mathbf{t} = \mathbf{l} \tag{9}$$

where $\oslash$ is the element-wise division, $\mathbf{I}$ is the unit matrix, the operator $Diag(\mathbf{v})$ is to construct a diagonal matrix using vector $\mathbf{v}$, and $\mathbf{D_d}$ are the Toeplitz matrices from the discrete gradient operators with forward difference. The main difference between our illumination map estimation method and that in [5] is the design of weight matrix $\mathbf{M}$. We adopted a simplified strategy which can yield similar results as in [5]. Other illumination decomposition techniques in Retinex-based methods can be borrowed here to find the weight matrix $\mathbf{W}$.

2.3 Camera Response Model

In our earlier work, we proposed a camera response model called Beta-Gamma Correction Model [16]. The BTF of our model is defined as

$$g(\mathbf{P}, k) = \beta \mathbf{P}^\gamma = e^{b(1-k^a)} \mathbf{P}^{(k^a)}. \tag{10}$$

where β and γ are two model parameters that can be calculated from camera parameters a, b and exposure ratio k. We assume that no information about the camera is provided and use a fixed camera parameters ($a = -0.3293, b = 1.1258$) that can fit most cameras.

2.4 Exposure Ratio Determination

In this subsection, we find the best exposure ratio so that the synthetic image is well-exposed in the regions where the original image under-exposed. First, we exclude the well-exposed pixels and obtain an image that is globally under-exposed. We simply extract the low illuminated pixels as

$$\mathbf{Q} = \{\mathbf{P}(x) | \mathbf{T}(x) < 0.5\}, \tag{11}$$

where $\mathbf{Q}$ contains only the under-exposed pixels.

The brightness of the images under different exposures changes significantly while the color is basically the same. Therefore, we only consider the brightness component when estimating k. The brightness component $\mathbf{B}$ is defined as the geometric mean of three channel:

$$\mathbf{B} := \sqrt[3]{\mathbf{Q}_r \circ \mathbf{Q}_g \circ \mathbf{Q}_b}, \tag{12}$$

where $\mathbf{Q}_r$, $\mathbf{Q}_g$ and $\mathbf{Q}_b$ are the red, green and blue channel of the input image $\mathbf{Q}$ respectively. We use the geometric mean instead of other definitions (*e.g.* arithmetic mean and weighted arithmetic mean) since it has the same BTF model parameters (β and γ) with all three color channels, as shown in Eq. 13.

$$\begin{aligned}\mathbf{B}' &:= \sqrt[3]{\mathbf{Q}'_r \circ \mathbf{Q}'_g \circ \mathbf{Q}'_b} \\ &= \sqrt[3]{(\beta\mathbf{Q}_r^{\gamma}) \circ (\beta\mathbf{Q}_g^{\gamma}) \circ (\beta\mathbf{Q}_b^{\gamma})} = \beta(\sqrt[3]{\mathbf{Q}_r \circ \mathbf{Q}_g \circ \mathbf{Q}_b})^{\gamma} \\ &= \beta\mathbf{B}^{\gamma}.\end{aligned} \tag{13}$$

The visibility of a well-exposed image is higher than that of an under-exposed image and it can provide a richer information for human. Thus, the optimal k should provide the largest amount of information. To measure the amount of information, we employ the image entropy which is defined as

$$\mathcal{H}(\mathbf{B}) = -\sum_{i=1}^{N} p_i \cdot \log_2 p_i, \tag{14}$$

where p_i is the i-th bin of the histogram of $\mathbf{B}$ which counts the number of data valued in $[\frac{i}{N}, \frac{i+1}{N})$ and N is the number of bins (N is often set to be 256). Finally, the optimal k is calculated by maximizing the image entropy of the enhancement brightness as

$$\hat{k} = \underset{k}{\operatorname{argmax}} \ \mathcal{H}(g(\mathbf{B}, k)). \tag{15}$$

The optimized k can be solved by one-dimensional minimizer. To improve the calculation efficiency, we resize the input image to 50×50 when optimizing k.

3 Experiments

To evaluate the performance of our method, we compare it with several state-of-the-art methods (AMSR [9], LIME [5], Dong [4] and NPE [14]) on hundreds of low-light images from five public datasets: VV[1], LIME-data [5], NPE [14] (NPE-data, NPE-ex1, NPE-ex2 and NPE-ex3), MEF [10], and IUS [8]. MEF and IUS are multi-exposure datasets, we select a low-light image from each multi-exposure set for evaluation.

3.1 Implementation Details

In our algorithm, μ is a parameter that controls the overall degree of enhancement. When $\mu = 0$, the resulting $\mathbf{R}$ is equal to $\mathbf{P}$, *i.e.*, no enhancement is performed. When $\mu = 1$, both the under-exposed pixels and well-exposed pixels are enhanced. When $\mu > 1$, pixels may get saturated and the resulting $\mathbf{R}$ suffers from detail loss. In order to perform enhancement while preserve the well-exposed regions, we set μ to 1/2.

[1] https://sites.google.com/site/vonikakis/datasets.

To maintain the fairness of the comparison, the parameters of our enhancement algorithm are fixed in all experiments: $\lambda = 1$, $\epsilon = 0.001$, $\mu = 1/2$, and the size of local window $\omega(x)$ is 5. The most time-consuming part of our algorithm is illumination map optimization. We employ the multi-resolution preconditioned conjugate gradient solver ($\mathcal{O}(N)$) to solve it efficiently. In order to further speedup our algorithm, we solve $\mathbf{T}$ using the down-sampled version of the input image and then up-sample the resulting $\mathbf{T}$ to the original size. No visual differences can be found in the enhanced results if we down-sample once, but the computational efficiency is greatly improved.

3.2 Contrast Distortion

As aforementioned, the image that only differ in exposures can be used as a reference for evaluating the accuracy of enhanced results. DRIM (Dynamic Range Independent Metric) [1] can measure the distortion of image contrast without the interference of change in image brightness. We use it to visualize the contrast difference between the enhancement result and the reference image.

As shown in Fig. 3, the proposed method obtains the most realistic results with the least distortion. The results of Dong suffer from severe contrast distortion. Although the details can be recovered by AMSR, the obvious loss of contrast makes the results look dim and unreal. By contrast, the results of LIME look a bit more vivid, but they suffer from amplification of invisible contrast. Figure 5 shows more examples for visual comparison.

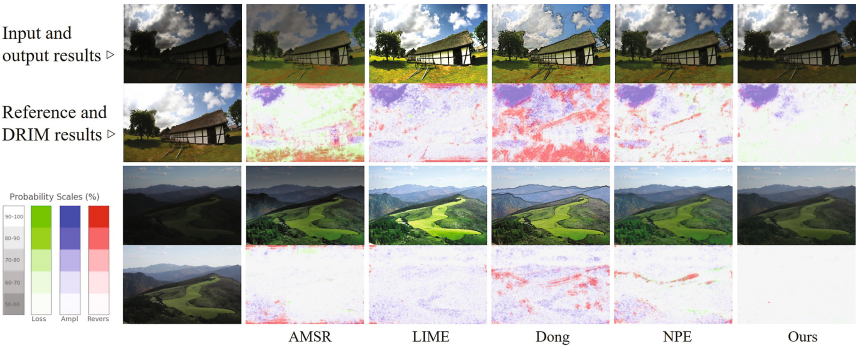

Fig. 3. Comparison of contrast distortion. The loss of visible contrast is marked in green, the amplification of invisible contrast is marked in blue, and the reversal of visible contrast is marked in red. Different shades of color represent different degrees of distortion. (Color figure online)

3.3 Lightness Distortion

We use lightness order error (LOE) to objectively measure the lightness distortion of enhanced results. LOE is defined as

$$LOE = \frac{1}{m}\sum_{x=1}^{m} RD(x) \tag{16}$$

where $RD(x)$ is the relative order difference of the lightness between the original image P and its enhanced version P' for pixel x, which is defined as follows:

$$RD(x) = \sum_{y=1}^{m} U\Big(\mathbf{L}(x), \mathbf{L}(y)\Big) \oplus U\Big(\mathbf{L}'(x), \mathbf{L}'(y)\Big), \tag{17}$$

where m is the pixel number, $\oplus$ stands for the exclusive-or operator, $\mathbf{L}(x)$ and $\mathbf{L}'(x)$ are the lightness component at location x of the input images and the enhanced images, respectively. The function $U(p, q)$ returns 1 if $p >= q$, 0 otherwise.

As suggested in [5,14], down-sampling is used to reduce the complexity of computing LOE. We noticed that LOE may change significantly when an image is down-sampled to different sizes since RD will increase as the pixel number m increases. Therefore, we down-sample all images to a fixed size. Specifically, we collect 100 rows and columns evenly to form a 100×100 down-sampled image. As shown in Table 1, our algorithm outperforms the others in all datasets. This means that our algorithm can maintain the naturalness of images well. We also provide a visualization of lightness distortion on two cases in Fig. 4, from which,

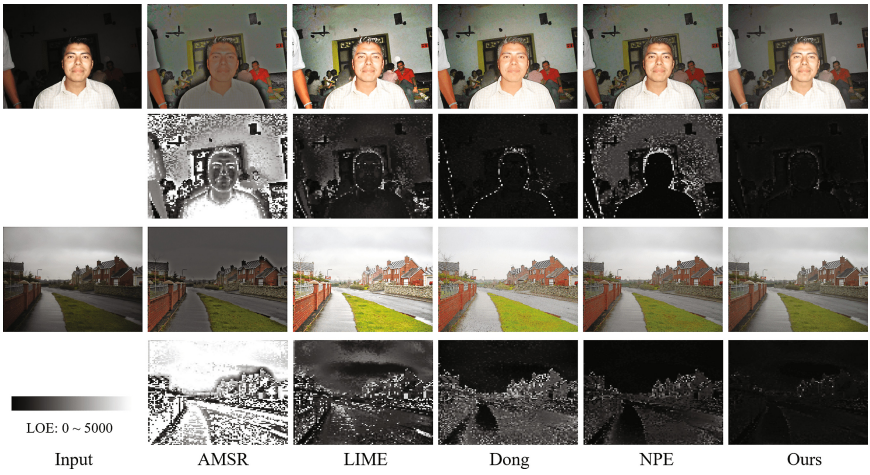

Fig. 4. Comparison of lightness distortion. The odd rows show the original image and the results of various enhancement methods, and the even rows show the visualization of each method's lightness distortion (RD).

Table 1. Qantitative measurement results of lighness distortion (LOE).

Methods	VV	LIME-data	NPE-data	NPE-ex1	NPE-ex2	NPE-ex3	MEF	IUS
AMSR	3892	3300	3298	2969	3212	3006	3264	4658
Dong	853	1244	1012	1426	1096	1466	1065	758
NPE	821	1471	646	841	776	1130	1158	559
LIME	1275	1324	1120	1322	1215	1319	1079	1306
Ours	**287**	**479**	**308**	**320**	**324**	**379**	**326**	**311**

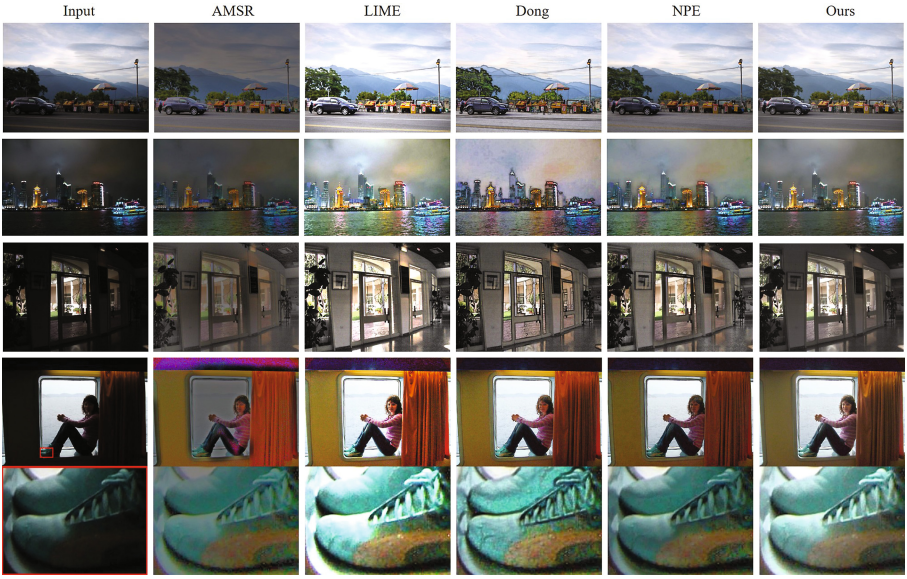

Fig. 5. Visual comparison among the competitors on different scenes.

we can find our results have the smallest lightness distortion. The results of AMSR lose the global lightness order and have the biggest lightness distortion. Although the results of LIME is visually pleasant, they also suffer from lightness distortion. The results of Dong and NPE can only retain the lightness order in the well-exposed regions.

4 Conclusion

In this paper, we propose an exposure fusion framework and an enhancement algorithm to provide an accurate contrast enhancement. Based on our framework, we solve three problems: (1) we borrow the illumination estimation techniques to obtain the weight matrix for image fusion. (2) we introduce our camera response model to synthesize multi-exposure images. (3) we find the best exposure ratio so that the synthetic image is well-exposed in the regions where the

original image under-exposed. The final enhancement result is obtained by fusing the input image and the synthetic image according to the weight matrix. The experimental results have revealed the advance of our method compared with several state-of-the-art alternatives. To encourage future works and allow more experimental verification and comparisons, we make the source code open. More testing results can be found on our project website: https://baidut.github.io/OpenCE/caip2017.html.

Acknowledgments. This work was supported by the grant of National Science Foundation of China (No.U1611461), Shenzhen Peacock Plan (20130408-183003656), and Science and Technology Planning Project of Guangdong Province, China (No. 2014B090910001 and No. 2014B010117007).

References

1. Aydin, T.O., Mantiuk, R., Myszkowski, K., Seidel, H.P.: Dynamic range independent image quality assessment. ACM Trans. Graph. (TOG) **27**(3), 69 (2008)
2. Beghdadi, A., Le Negrate, A.: Contrast enhancement technique based on local detection of edges. Comput. Vis. Graph. Image Process. **46**(2), 162–174 (1989)
3. Chen, S.D., Ramli, A.R.: Minimum mean brightness error bi-histogram equalization in contrast enhancement. IEEE Trans. Consum. Electron. **49**(4), 1310–1319 (2003)
4. Dong, X., Wang, G., Pang, Y., Li, W., Wen, J., Meng, W., Lu, Y.: Fast efficient algorithm for enhancement of low lighting video. In: 2011 IEEE International Conference on Multimedia and Expo, pp. 1–6. IEEE (2011)
5. Guo, X.: Lime: a method for low-light image enhancement. In: Proceedings of the 2016 ACM on Multimedia Conference, pp. 87–91. ACM (2016)
6. Ibrahim, H., Kong, N.S.P.: Brightness preserving dynamic histogram equalization for image contrast enhancement. IEEE Trans. Consum. Electron. **53**(4), 1752–1758 (2007)
7. 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 Trans. Image Process. **6**(7), 965–976 (1997)
8. Karaduzovic-Hadziabdic, K., Telalovic, J.H., Mantiuk, R.: Subjective and objective evaluation of multi-exposure high dynamic range image deghosting methods (2016)
9. Lee, C.H., Shih, J.L., Lien, C.C., Han, C.C.: Adaptive multiscale retinex for image contrast enhancement. In: 2013 International Conference on Signal-Image Technology & Internet-Based Systems (SITIS), pp. 43–50. IEEE (2013)
10. Ma, K., Zeng, K., Wang, Z.: Perceptual quality assessment for multi-exposure image fusion. IEEE Trans. Image Process. **24**(11), 3345–3356 (2015)
11. Peli, E.: Contrast in complex images. JOSA A **7**(10), 2032–2040 (1990)
12. Reza, A.M.: Realization of the contrast limited adaptive histogram equalization (clahe) for real-time image enhancement. J. VLSI Signal Process. Syst. Signal Image Video Technol. **38**(1), 35–44 (2004)
13. Wang, C., Ye, Z.: Brightness preserving histogram equalization with maximum entropy: a variational perspective. IEEE Trans. Consum. Electron. **51**(4), 1326–1334 (2005)

14. Wang, S., Zheng, J., Hu, H.M., Li, B.: Naturalness preserved enhancement algorithm for non-uniform illumination images. IEEE Trans. Image Process. **22**(9), 3538–3548 (2013)
15. Xu, L., Yan, Q., Xia, Y., Jia, J.: Structure extraction from texture via relative total variation. ACM Trans. Graph. (TOG) **31**(6), 139 (2012)
16. Ying, Z., Li, G., Ren, Y., Wang, R., Wang, W.: A new low-light image enhancement algorithm using camera response model, manuscript submitted for publication (2017)