A Bio-Inspired Multi-Exposure Fusion Framework for Low-light Image Enhancement

Zhengqiang Ying, Student Member, IEEE, Ge Li, Member, IEEE, and Wen Gao, Fellow, IEEE

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. Inspired by human visual system, we design a multi-exposure fusion framework for low-light image enhancement. Based on the framework, we propose a dual-exposure fusion algorithm to provide an accurate contrast and lightness 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 is under-exposed. Finally, the enhanced result is obtained by fusing the input image and the synthetic image according to the weight matrix. Experiments show that our method can obtain results with less contrast and lightness distortion compared to that of several state-of-the-art methods.

Index Terms—Image enhancement, contrast enhancement, exposure compensation, exposure fusion.

I. INTRODUCTION

With the development of photographic techniques, the image quality is greatly improved in both resolution and bit-depth. However, images captured by standard imaging devices often suffer from low visibility in non-uniform illuminated environments such as back lighting, nighttime and low-light indoor scene. Those images may lose information in under-exposed regions, making the image content invisible to human eyes. Since the camera dynamic range is limited, if we increase camera exposure to reveal the information of under-exposed regions, the well-exposed regions will be over-exposed or even saturated. To address the problem, many image enhancement techniques have been proposed including histogram-based methods [1], [2], [3], [4], [5], [6], [7], [8], [9], [10], [11], [12], [13]. Retinex-based methods [14], [15], [16], [17], [18], [19], [20]. Logarithmic Image Processing methods [21], [22], and filtering-based methods [23], [24], [25], [26], [27]. Although some methods can obtain results with good subjective quality, those results may not accurately reflect the true lightness and contrast of the scene. So, accurate light and contrast enhancement based on a single image is still a challenging problem.

With a set of different exposure images in the same scene, High Dynamic Range (HDR) techniques can synthesize images that are close to the perceived scene. We know that the camera and the human eye have a lot of similarities, then, why can we perceive an image that is well-exposed everywhere while the camera cannot? The reason lies in the post-processing in our brains that have an image fusion mechanism similar to HDR technique [28]. Human eye exposure changes with its focus point, resulting in multi-exposure image set which is then sent to the brain. Although each one of those images suffers from under-exposed or over-exposed problem in some regions, our brain can fuse these images into an image that is free from under- and over-exposed problems, as shown in Fig. 1.

Can we introduce this fusion mechanism of human visual system (HVS) to help build an accurate image enhancement algorithm? Although many exposure fusion techniques have been proposed in the domain of HDR, the additional images taken with different exposures are often not available for the low-light enhancement problem. Fortunately, those images are highly correlated. The mapping function between two images that only differ in exposure is called Brightness Transform Function (BTF). Therefore, we can first use BTF to generate a series of multi-exposure images and then fuse those images to obtain the enhanced result.

In this paper, we propose a multi-exposure fusion framework inspired by the HVS. There are two stages in our framework: Eye Exposure Adjustment and Eye Exposure Adjustment. The first stage simulates the human eye to adjust the exposure, generating an multi-exposure image set. The second stage simulates the human brain to fuse the generated
images into the final enhanced result. Based on our framework, we propose a dual-exposure fusion method. Specifically, we first employ the illumination estimation techniques to build the weight matrix for image fusion. Then we derive our camera response model based on observation. Next, we find the optimal exposure for our camera response model to generate the synthetic image that is well-exposed in the regions where the original image is under-exposed. Finally, we obtain the enhanced results by fusing the input image with the synthetic image using the weight matrix. Experiments on five challenging datasets are conducted to reveal the advantages of our method in comparison with other state-of-the-art methods.

II. RELATED WORK

In general, image enhancement techniques can improve the subjective visual quality of input images and support the extraction of valuable information for some computer vision techniques [29], [16]. Low-light image enhancement, as one of enhancement techniques, can reveal the information of the under-exposed regions in an image. Broadly speaking, existing low-light image enhancement techniques can be divided into two major categories: global enhancement and local enhancement.

A. Global Enhancement Algorithms

Global enhancement performs same processing on all image pixels regardless of their spatial distribution. Linear amplifying is a simple and straightforward global enhancement method. However, bright regions might be saturated after linear amplifying, causing some detail loss to the enhanced results. To avoid the problem, some image enhancement methods adopt non-linear monotonic functions (e.g. power-law [30], logarithm [31] and gamma function [32]) to perform enhancements. As another way to avoid saturation, histogram equalization (HE) [33] can improve the contrast effectively and became a widely-used technique. Many extensions of HE are proposed to take some restrictions into account such as brightness preservation [29], [34], [10] and contrast limitation [35]. By extending the notion of the histogram, some algorithms take spatial image features into consideration to further improve the performance [3], [6], [7]. However, global enhancement may suffer from detail loss in some local areas because a global processing can not ensure all local areas be well enhanced.

B. Local Enhancement Algorithms

By making use of spatial information directly, local enhancement can achieve better results and become the mainstream of recent techniques. Local histogram equalization [36], [37] adopts the sliding window strategy to perform HE locally. Based on the observation that the inverted low-light images are closed to hazy images, dehazing techniques are borrowed to solve low-light image enhancement in some methods [38], [39]. However, the basic models of above methods are lacking in physical explanation [14]. To provide a physical meaningful model for image enhancement, Retinex theory assumes that the amount of light reaching observers can be decomposed into two parts: illumination and scene reflection. Most Retinex-based methods get enhanced results by removing the illumination part [18], [20], [15] while others [40], [16], [14] keep a portion of the illumination to preserve naturalness. Fu et al. [41] adjust the illumination components by fusing it with two enhanced illumination. As far as we know, there is no multi-exposure fusion method for this task since lowlight enhancement problem usually takes a single image as input.

III. MULTI-EXPOSURE FUSION FRAMEWORK

Our framework mainly consists of four main components: the first component, named Multi-Exposure Sampler, determines how many images are required and the exposure ratio of each image to be fused; the second component, named Multi-Exposure Generator, use a camera response model and the specified exposure ratio to synthetic multi-exposure images; the third component, named Multi-Exposure Evaluator, determines the weight map of each image when fusing; the last component, named Multi-Exposure Combiner, is to fuse the generated images to the final enhanced result based on the weight maps. In this section, we introduce them one by one.

A. Multi-Exposure Sampler

Before we generate multi-exposure images, we need to determine how many images are required and their exposure ratios. Since some images in the multi-exposure set cannot
provide additional information, taking these images into consideration is a waste of computation resources and may even deteriorate the fused result. A good sampler can use as few images as possible to reveal all the information in a scene by choosing appropriate exposure ratios. The output of the sampler is a set of exposure ratios \( \{k_1, k_2, \ldots, k_N\} \) where \( N \) is the number of the images to be generated.

### B. Multi-Exposure Generator

As aforementioned, images taken with different exposures are correlated. Multi-Exposure Generator maps the input image into multi-exposure images according to the specified exposure ratio set. The key part of the Multi-Exposure Generator is the camera response model used to find an appropriate BTF for mapping. Given an exposure ratio \( k_i \) and a BTF \( g \), we can map the input image \( P \) to the \( i \)-th image in the exposure set as

\[
P_i = g(P, k_i).
\]  

### C. Multi-Exposure Evaluator

To estimate the wellness of each pixel in the generated images. The Multi-Exposure Evaluator takes an image and outputs a weight matrix that indicates the wellness of each pixel. The weight matrix is nonuniform for all pixels: the well-exposed pixels are enhanced. When \( \mu > 1 \), pixels may get saturated and the resulting \( R \) suffers from detail loss. As shown in Fig. 3. In order to perform enhancement while preserving the well-exposed regions, we set \( \mu \) to 0.5. The scene illumination map \( T \) is estimated by solving an optimization problem.

1) **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:

\[
L(x) = \max_{c \in \{R,G,B\}} P_c(x)
\]  

for each individual pixel \( x \). Ideal illumination should have 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 [13], we refine \( T \) by solving the following optimization equation:

\[
\min_T ||T - L||^2_2 + \lambda ||M \circ \nabla T||_1,
\]  

where \( ||*||_2 \) and \( ||*||_1 \) are the \( \ell_2 \) and \( \ell_1 \) norm, respectively. The first order derivative filter \( \nabla \) contains \( \nabla_h T \) (horizontal) and \( \nabla_v T \) (vertical). \( M \) is the weight matrix and \( \lambda \) 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 [44]. 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

\[ M_d(x) = \frac{1}{|\sum_{y \in \omega(x)} \nabla_d L(y)| + \epsilon}, \quad d \in \{h, v\}, \tag{8} \]

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

2) *Closed-Form Solution:* To reduce the complexity, we approximate Eq. 7 as in [14]:

\[
\min_T \sum_x \left( (T(x) - L(x))^2 + \lambda \sum_{d \in \{h, v\}} \frac{M_d(x)(\nabla_d T(x))^2}{|\nabla_d L(x)| + \epsilon} \right). \tag{9}
\]

As can be seen, the problem now only involves quadratic terms. Let \(m_d, l, t\) and \(\nabla_d l\) denote the vectorized version of \(M_d, L, T\) and \(\nabla_d L\) respectively. Then the solution can be directly obtained by solving the following linear function.

\[
(I + \lambda \sum_{d \in \{h, v\}} (D_d^T \text{Diag}(m_d \odot (|\nabla_d l| + \epsilon))D_d)t = 1 \tag{10}
\]

where \(\odot\) is the element-wise division, \(I\) is the unit matrix, the operator \(\text{Diag}(v)\) is to construct a diagonal matrix using vector \(v\), and \(D_d\) are the Toeplitz matrices from the discrete gradient operators with forward difference.

The main difference between our illumination map estimation method and that in [14] is the design of weight matrix \(M\). We adopt a simplified strategy which can yield similar results as in [14]. As shown in Fig. 4, although the illumination map in [14] is sharper than ours, our method is more time-saving while the two enhanced results show no significant visual difference.

### B. Dual-Exposure Generator

In this section, we present a camera response model to implement the Multi-Exposure Generator. A camera response model consists of two parts: Camera Response Function (CRF) model and BTF model. The parameters of CRF model is determined only by camera while that of BTF model is determined by camera and exposure ratio. In this subsection, we first propose our BTF model based on the observation of two different exposure images. Then we derive the corresponding CRF model by solving the compararative equation. Finally, we discuss how to determine the model parameters and present the final form of \(g\).

1) *BTF Estimation:* To estimate the BTF \(g\), we select a pair of images \(P_0\) and \(P_1\) that differ only in exposure. Then we plot their histograms of each color channel, as shown in Fig. 5. Noticing that the histograms of the under-exposed image mainly concentrate in low-brightness area, if we perform linear amplification of pixel values before traditional gamma correction, then the resulting image will be very close to the real well-exposed image. Therefore, we can use a two-parameter function to describe the BTF model as

\[
P_1 = g(P_0, k) = \beta P_0^\gamma, \tag{11}
\]

where \(\beta\) and \(\gamma\) are parameters in our BTF model related to exposure ratio \(k\). The observation also shows that different color channels have approximately same model parameters. The underlying reason is that the response curves of different color channels are approximately identical for general cameras.

2) *CRF Estimation:* In our BTF model, \(\beta\) and \(\gamma\) are determined by the camera parameters and exposure ratio \(k\). To find their relationship, we need to obtain the corresponding CRF model. The CRF model can be derived by solving the following compararative equation (plug \(g = \beta f^\gamma\) to \(f(kE) = g(f(E))\)):

\[
f(kE) = \beta f(E)^\gamma. \tag{12}
\]

The closed-form solution of \(f\) is provided in as follows (see Appendix for detail):

\[
f(E) = \begin{cases} e^{b(1-E^a)}, & \text{if } \gamma \neq 1, \\ E^c, & \text{if } \gamma = 1. \end{cases} \tag{13}
\]

where \(a\) and \(b\) are model parameters in the case of \(\gamma \neq 1\):

\[
a = \log_k \gamma, \quad b = \frac{\ln \beta}{1 - \gamma}; \tag{14}
\]

And \(c\) is a model parameter in the case of \(\gamma = 1\):

\[
c = \log_k \beta. \tag{15}
\]
geometric mean of three channel: \( k \)

Therefore, we only consider the brightness component when changes significantly while the color is basically the same.

where \( Q \) instead of other definitions (e.g. the input image \( Q \) 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 \( B \) is defined as the geometric mean of three channel:

\[
B := \sqrt[3]{Q_r \circ Q_g \circ Q_b},
\]

where \( Q_r \), \( Q_g \) and \( Q_b \) are the red, green and blue channel of the input image \( Q \) 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 (\( \beta \) and \( \gamma \)) with all three color channels, as shown in Eq. [19]

\[
B' := \sqrt[3]{Q_r' \circ Q_g' \circ Q_b'}
\]

where \( Q_r' \), \( Q_g' \) and \( Q_b' \) are the red, green and blue channel of the input image \( Q' \) 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 (\( \beta \) and \( \gamma \)) with all three color channels, as shown in Eq. [19]

\[
B' := \sqrt[3]{(\beta Q_r' \circ (\beta Q_g') \circ (\beta Q_b') } = \beta B^\gamma.
\]  

The visibility of a well-exposed image is higher than that of an under/over-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

\[
\mathcal{H}(B) = - \sum_{i=1}^{N} p_i \cdot \log_2 p_i,
\]

where \( p_i \) is the \( i \)-th bin of the histogram of \( B \) which counts the number of data valued in \( \left[ \frac{i}{N}, \frac{i+1}{N} \right) \) and \( N \) is the number of bins (\( N \) is often set to be 256). As shown in Fig. 6 the image entropy of a well-exposed image is higher than that of an under/over-exposed image. Therefore, it is reasonable to use the entropy to find the optimal exposure ratio. The optimal exposure ratio \( \hat{k} \) is calculated by maximizing the image entropy of the enhancement brightness as

\[
\hat{k} = \argmax_k \mathcal{H}(g(B, k)).
\]

Since the image entropy increases first and then decreases with the increase of the exposure ratio, \( \hat{k} \) can be solved by one-dimensional minimizer. To improve the calculation efficiency, we resize the input image to 50 \( \times \) 50 when optimizing \( k \).

V. Experiments

To evaluate the performance of our method, we compare it with several state-of-the-art methods, including Multi Scale Retinex with Color Restoration (MSRCR) [45], Naturalness Preserved Enhancement algorithm (NPE) [16], dehazing based method (Dong) [38], Multi-deviation Fusion method (MF) [41], Illumination Estimation based method (LIME) [14] and Simultaneous Reflection and Illumination Estimation (SRIE) [46]. We test those methods on hundreds of low-light images from five public datasets: VV [47], LIME-data [14], NPE-data and its extension (NPE, NPE-ex1, NPE-ex2 and NPE-ex3) [16], DICM [3], and MEF [48]. The datasets are briefly introduced as follows:
Fig. 7. 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 I

| Dataset        | VV     | LIME-data | NPE-data | NPE-ex1 | NPE-ex2 | NPE-ex3 | DICM     | MEF     |
|----------------|--------|-----------|----------|---------|---------|---------|----------|---------|
| MSRCR          | 2727.7 | 1835.5    | 1889.7   | 1870.3  | 1944.7  | 1776.3  | 1795.3   | 1686.2  |
| Dong           | 853.35 | 1244      | 1012     | 1426.1  | 1096.3  | 1466.2  | 1180     | 1065.4  |
| NPE            | 820.93 | 1471.3    | 646.34   | 840.83  | 775.82  | 1130    | 662.29   | 1158.2  |
| LIME           | 1274.7 | 1323.8    | 1119.6   | 1321.9  | 1215.4  | 1319.1  | 1260.8   | 1079.4  |
| MF             | 470.93 | 629.82    | 488.07   | 851.87  | 541.85  | 749.72  | 667.45   | 525.95  |
| SRIE           | 551.39 | 823.61    | 533.24   | 653.05  | 564.49  | 760.76  | 623.32   | 754.2   |
| Ours           | 287.22 | 478.57    | 308.12   | 319.93  | 323.72  | 378.65  | 351.82   | 325.86  |

**VV** This dataset is collected by Vassilios Vonikakis in his daily life to provide the most challenging cases for enhancement. Each image in the dataset has a part that is correctly exposed and a part that is severely under/over-exposed. A good enhancement algorithm should enhance the under/over-exposed regions while not affect the correctly exposed one.

**LIME-data** This dataset contains 10 low-light images used in [4].

**NPE** This dataset contains 85 low-light images downloaded from Internet. NPE-data contains 8 outdoor nature scene images which are used in [16]. NPE-ex1, NPE-ex2 and NPE-ex3 are three supplementary datasets including cloudy daytime, daybreak, nightfall and nighttime scenes.

**DICM** It contains 69 captured images from commercial digital cameras collected by [3].

**MEF** This dataset was provided by [48]. It contains 17 high-quality image sequences including natural sceneries, indoor and outdoor views and man-made architectures. Each image sequence has several multi-exposure images, we select one of poor-exposed images as input to perform evaluation.

In order to maintain the fairness of the comparison, all the codes are in Matlab and all the experiments are conducted on a PC running Windows 10 OS with 64G RAM and 3.4GHz CPU (GPU acceleration is not used). 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 ($O(N)$) to solve it efficiently.

### A. Lightness Distortion

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

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

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:

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

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

As suggested in [14], [16], down-sampling is used to reduce the complexity of computing LOE. We notice 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 \( 100 \times 100 \) down-sampled image.

As shown in Table I, 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. 7, from which, we can find our results have the smallest lightness distortion. The results of MSRCR lose the global lightness order and suffer from severe lightness distortion. Although the results of LIME is visually pleasant, they are full of lightness distortion. The results of Dong, NPE, MF and SRIE can only retain the lightness order in the well-exposed regions.

### B. 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) [49] 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 enhanced result and the reference image. As shown in Fig. 8, the proposed method obtains the most realistic results with the least distortion.

### C. Visual Information Distortion

To measure the distortion of visual information, we employ Visual Information Fidelity (VIF) [50] in reverse mode. As a full reference image quality assessment index, VIF models the quality assessment problem as a information fidelity criterion that quantifies the mutual information between the reference image \( C \) and the distorted image \( F \) relative to the information of \( C \) extracted by the HVS. VIF measure is given by

\[
VIF = \frac{I(C; F)}{I(C; E)},
\]

where \( E \) is the image that the HVS perceives. The mutual information \( I(C; F) \) and \( I(C; E) \) represent the information that could be extracted by the brain in the reference and the test images respectively.

Like most of full reference image quality assessment methods, VIF were designed for and tested on degraded images. The normal version of VIF treats the original image as the reference image and the image outputted by algorithm as the degraded image. For image enhancement problem, however, the original image is the degraded one. Therefore, we employ VIF in reverse mode by specifying the enhanced version of the image as the reference and the original image as the degraded image. VIF provides consistently high value of correlation between subjective MOS (Mean Opinion Score) and its scores, as...
shown in [51]. Besides, it is suitable for automating evaluation process of nonlinear image enhancement algorithms [52]. As shown in Table II, our algorithm outperforms the others in all datasets. This means that our algorithm can maintain the visual information of images well.

D. Time Cost

Fig. 10 gives the comparison among different methods in terms of time cost. Although SRIE and NPE produce small distortion, they are quite time-consuming. Our method achieves the smallest distortion than others with an acceptable time cost.

E. Subjective Evaluation

Fig. 9 shows more examples for visual comparison. Although the color correction post-processing in MSRCR can handle the color cast in some cases (e.g. underwater and hazy images), it may make the results look whitish. Besides, in some dark show halo artifacts around sharp edges. The results of MSRCR shows severe halo artifacts around sharp edges and obvious noise in very dark areas (see the two astronauts) and suffer from detail loss in some bright areas (see the Christmas hat). The results of Dong is noisy and full of bold edges making it look like exaggerated art paintings. The results of LIME is so bright that many bright areas are saturated. Also, the noise in dark areas are amplified and de-noising methods are therefore required to obtain better results. MF may introduce color over-enhancement (see the ground beneath the astronauts’ feet) and SRIE may produce slight halo effects in some edges (see the Christmas hat).

VI. LIMITATION AND FUTURE WORK

Fig. 11 shows a failure case of our technique that the hair of the man turns to be grey because of over-enhancement. This is due to the dark area behind his head blending with his black hair. As shown in Fig. 11 (c), the hair is mistaken as the dark background in the estimated illumination map and therefore is enhanced along with the background. Such mistake is a result of the existing illumination map estimation techniques. This highlights a direction for future work. To avoid the over-enhancement due to the ignorance of the scene content, semantic understanding is required. With further refinement, we might employ the deep learning techniques to estimate the illumination map.

Besides, we only use two images to obtain the enhanced result. The over-exposure problem is remain unsolved. Images with smaller exposures than the input image should be considered in our framework to obtain a better result. We will address this problem as future work.
VII. 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 enhanced 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 to reproduce the experiment in the paper. We also provide the results of other competitors to facilitate the validation of quality measures.

APPENDIX

DERIVATION OF CAMERA RESPONSE MODEL

The CRF model can be derived by solving the following compararative equation (plug $g = \beta f^\gamma$ to $f(kE) = g(f(E), k)$):

$$f(kE) = \beta f(E)^\gamma.$$  
(25)

Since $\beta$ is positive, we can take the logarithm of both sides of Eq. 12:

$$\ln(f(kE)) = \ln(\beta) + \gamma \ln(f(E)).$$  
(26)

Differentiate with respect to $E$, we get

$$k \frac{f'(kE)}{f(kE)} = \gamma f'(E)f(E).$$  
(27)

The power function $f'(kE)/f(kE) = C_1 E^{a-1}$ where $a = \log_k \gamma$ can be a solution of Eq. 21. Then $f$ can be solved by integration as

$$f(E) = \begin{cases} e^{C_1 \frac{E^a}{a} + C_2}, & \text{if } \gamma \neq 1, \\ C_3 E^{C_2}, & \text{if } \gamma = 1. \end{cases}$$  
(28)

The constants $C_1$, $C_2$ and $C_3$ can be determined by the restriction of $f(1) = 1$ and $f(k) = \beta$ (plug $f(1) = 1$ to Eq. 12) and the resulting $f$ is obtained as follows:

$$f(E) = \begin{cases} e^{(b-1-E^a)}, & \text{if } \gamma \neq 1, \\ E^c, & \text{if } \gamma = 1, \end{cases}$$  
(29)

where $a$ and $b$ are model parameters in the case of $\gamma \neq 1$:

$$a = \log_k \gamma, \quad b = \ln \frac{\beta}{1 - \gamma};$$  
(30)

And $c$ is model parameter in the case of $\gamma = 1$:

$$c = \log_k \beta.$$  
(31)

REFERENCES

[1] H. Xu, G. Zhai, X. Wu, and X. Yang, “Generalized equalization model for image enhancement,” IEEE Transactions on Multimedia, vol. 16, no. 1, pp. 68–82, 2014.

[2] T. Celik, “Spatial entropy-based global and local image contrast enhancement,” IEEE Transactions on Image Processing, vol. 23, no. 12, pp. 5298–5308, 2014.

[3] C. Lee, C. Lee, and C.-S. Kim, “Contrast enhancement based on layered difference representation,” in Image Processing (ICIP), 2012 19th IEEE International Conference on. IEEE, 2012, pp. 965–968.

[4] K. Arya, M. Pattanaik et al., “Histogram statistics based variance controlled adaptive threshold in anisotropic diffusion for low contrast image enhancement,” Signal Processing, vol. 93, no. 6, pp. 1684–1693, 2013.

[5] C. Lee, C. Lee, Y.-Y. Lee, and C.-S. Kim, “Power-constrained contrast enhancement for emissive displays based on histogram equalization,” IEEE transactions on image processing, vol. 21, no. 1, pp. 80–93, 2012.

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

[7] T. Arici, S. Dikbas, and Y. Aftunbasak, “A histogram modification framework and its application for image contrast enhancement,” IEEE Transactions on Image Processing, vol. 18, no. 9, pp. 1921–1935, 2009.

[8] D. Coltuc, P. Bolon, and J.-M. Chassery, “Exact histogram specification,” IEEE Transactions on Image Processing, vol. 15, no. 5, pp. 1143–1152, 2006.

[9] S.-D. Chen and A. R. Ramli, “Contrast enhancement using recursive mean-separate histogram equalization for scalable brightness preservation,” IEEE Transactions on Consumer Electronics, vol. 49, no. 4, pp. 1301–1309, 2003.

[10] K. Muller, M. P. Braeuning, and S. M. Pizer, “Contrast limited adaptive histogram equalization algorithm for non-uniform illumination images,” IEEE Transactions on Image Processing, vol. 23, no. 8, pp. 3381–3396, 2014.

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

[12] E. D. Pisano, S. Zong, B. M. Heineninger, M. DeLuca, R. E. Johnston, K. Muller, M. P. Braeuning, and S. M. Pizer, “Contrast limited adaptive histogram equalization image processing to improve the detection of simulated spiculations in dense mammograms,” Journal of Digital imaging, vol. 11, no. 4, pp. 193–200, 1998.

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

[14] X. Guo, Y. Li, and H. Ling, “Lime: Low-light image enhancement via illumination map estimation,” IEEE Transactions on Image Processing, vol. 26, no. 2, pp. 982–993, 2017.

[15] L. Wang, L. Xiao, H. Liu, and Z. Wei, “Variational bayesian method for retinex,” IEEE Transactions on Image Processing, vol. 23, no. 8, pp. 3538–3548, 2013.

[16] S. Wang, J. Zheng, H.-M. Hu, and B. Li, “Naturalness preserved enhancement algorithm for non-uniform illumination images,” IEEE Transactions on Image Processing, vol. 22, no. 9, pp. 3538–3548, 2013.

[17] R. Kimmel, M. Elad, D. Shaked, R. Keshet, and I. Sobel, “A variational framework for retinex,” International Journal of computer vision, vol. 16, no. 1, pp. 7–23, 2003.

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