CODEN: combined optimization-based decomposition and learning-based enhancement network for Retinex-based brightness and contrast enhancement

SANGJAE AHN,[1] JOONGCHOL SHIN,[1] HEUNSEUNG LIM,[1] JAEHEE LEE,[1] AND JOONKI PAIK[1,2,*] (iD)

[1]*Department of Image, Chung-Ang University, 84 Heuseok-Ro, Dongjak-Gu, Seoul, Republic of Korea*
[2]*Department of Artificial Intelligence, Chung-Ang University, 84 Heuseok-Ro, Dongjak-Gu, Seoul, Republic of Korea*
*paikj@cau.ac.kr

Abstract: In this paper, we present a novel low-light image enhancement method by combining optimization-based decomposition and enhancement network for simultaneously enhancing brightness and contrast. The proposed method works in two steps including *Retinex decomposition* and *illumination enhancement*, and can be trained in an end-to-end manner. The first step separates the low-light image into illumination and reflectance components based on the Retinex model. Specifically, it performs model-based optimization followed by learning for edge-preserved illumination smoothing and detail-preserved reflectance denoising. In the second step, the illumination output from the first step, together with its gamma corrected and histogram equalized versions, serves as input to illumination enhancement network (IEN) including residual squeeze and excitation blocks (RSEBs). Extensive experiments prove that our method shows better performance compared with state-of-the-art low-light enhancement methods in the sense of both objective and subjective measures.

1. Introduction

Brightness and contrast of digital images serve as a key information for various computer vision techniques including object detection, image recognition, visual surveillance, and image segmentation [1–6]. However, most imaging systems cannot avoid undesired artifacts such as low-light, low contrast, motion blur and haze during the image acquisition process, and those artifacts degrades the performance of vision tasks [7,8]. Among various image enhancement methods, low-light image enhancement improves the brightness and contrast of images taken under the low-light environment [9–11].

Existing contrast enhancement approaches include contextual and variational contrast enhancement (CVC) [12] and layed difference representation (LDR) [13] methods. They used the difference in brightness between adjacent pixels to create a smoothed histogram. Also, there are multi-branch low-light enhancement network (MBLLEN) and Retinex-Net methods for brightness enhancement using deep learning methods [14]. The MBLLEN first extracted feature maps of various levels, and then combines enhanced feature maps [14]. The Retinex-Net restored low-light images using decomposition network to separate illumination and reflectance, enhancement network for brightness enhancement, and additional denoising operation [15,16]. Although these methods successfully enhanced the brightness of low-light images using deep neural networks trained end-to-end, there is a performance limitation to simultaneously improving the brightness and contrast. The original Retinex theory served a fundamental basis of decomposing an image into the illuminance and reflectance components, but there was no practically optimal methods for the decomposition for a long time. Although the Retinex-Net first presented a deep learning-based

decomposition, it provided neither in-depth research on the enhancement network architecture nor discussion of the training datase. To solve this problems, we present a novel method for simultaneously enhancing brightness and contrast based on the Retinex model.

The proposed method consists of two steps: *decomposition* and *enhancement*. The decomposition step separates the low-light image into illumination and reflectance components based on the Retinex model. Various studies of Retinex decomposition have been conducted. The model-based optimization approach is flexible in handling various regularized restoration problems at the cost of increased computational complexity [17]. On the other hand, the learning-based approach has an advantage of fast inference speed and end-to-end training [17]. For that reason, the proposed method takes the optimization approach to decompose the illumination and reflectance components, and then uses learning-based approach to enhance the low-light illumination component.

Even if the estimation of the illumination is successful, further visual enhancement is not possible when the input image has a low brightness and/or contrast. To solve the under-enhancement problem with limited brightness and contrast in the dehazing application, Ren et al. proposed gated fusion using multiple derived inputs [18]. Originated from that idea, in the second step of the proposed method, illumination enhancement network (IEN) takes a set of three derived inputs including: i) the illumination component from the first decomposition step, ii) the contrast enhanced (or histogram equalized) illumination, and iii) the gamma corrected illumination. The IEN improves the brightness and contrast by rearranging the importance of feature maps through residual squeeze and excitation block (RSEB) [19].

An example of low-light image enhancement results are shown in Fig. 1. Compared with two state-of-the-art methods, the proposed method better improves both brightness and contrast. In addition, the noise is well removed because the reflectance is estimated by combining the optimization and learning steps. On the other hand, low-light image enhancement (LIME) and Hao's methods cannot completely remove the noise because they used only optimization step [20,21].
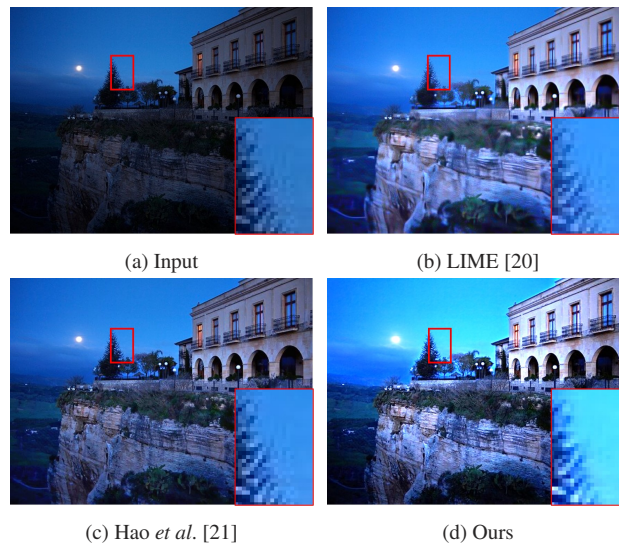

(a) Input

(b) LIME [20]

(c) Hao *et al.* [21]

(d) Ours

Fig. 1. Visual comparison of low-light image enhancement methods.

Major contributions of the proposed method are threefold as:

1 We proposed a combined optimization and deep learning framework for simultaneous bright and contrast enhancement. The proposed low-light image enhancement method can successfully decompose edge-preserved illumination and detail-preserved reflectance components.

2 We created new derived inputs including the decomposed illumination, its gamma-corrected and histogram-equalized versions to overcome the limitation of low brightness image enhancement.

3 We proposed illumination enhancement net (IEN) including RSEB to improve the brightness and contrast of the resulting image by readjusting the importance of feature maps.

The major advantage of an optimization-based image enhancement method is to provide the optimal solution of a user-defined objective function. During the past decades, various objective functions for image enhancement have been proposed and successfully employed in most image enhancement methods. On the other hand, enhanced images should satisfy human subjective criterion that is based on the human cognitive knowledge through experiencing many images. In this context, learning-based methods can provide subjectively satisfying result if a training set of images is well-prepared. Different advantages of optimization-based and learning-based methods cannot be easily combined due to the nature of each method. The proposed method successfully combined both advantages, and proved their performance using various ablation studies.

2. Related works

2.1. Conventional methods

Early works for low-light image enhancement used histogram equalization. Various histogram-based enhancement methods were proposed in the literature [12,13,22,23].

The Retinex theory assumes that an image can be decomposed into reflectance and illumination, and the amount of illumination only affects the illumination component. Early Retinex model-based methods include single-scale Retinex (SSR) [24] and multiscale Retinex (MSR) [25]. Adaptive multiscale Retinex (AMSR) adaptively improved the brightness of the input image by using the weight values associated with each SSR output image [24–26]. Wang et al. proposed naturalness preserved enhancement (NPE) method for non-uniformly illuminatied images to preserve naturalness [27]. They also proposed an improved method to mimic the human visual system that perceives colors [28].

In addition to histogram equalization- and Retinex-based methods, various types of image filtering were used for low-light image enhancement [29–32].

2.2. Model-based optimization methods

Various model-based optimization methods were proposed in the field of low-light image enhancement. Cai et al. used Retinex-based joint intrinsic-extrinsic prior (JIEP) model to decompose the illumination and reflectance components [33]. Fu et al. proposed a new probabilistic method to simultaneously estimate the illumination and reflectance components in the linear domain [34]. They also proposed simultaneous reflectance and illumination estimation (SRIE) method using a weighted variational model [35]. Lee et al. solved a constrained optimization problem for Retinex decomposition using the gray-level differences [13]. Ren et al. proposed a joint enhancement and denoising (JED) method that sequentially performs low-light enhancement and noise reduction [36]. They also proposed low-light image enhancement using camera response model (LECARM) algorithm using camera characteristics-based optimization [37]. Guo et al. proposed low-light image enhancement via illumination map estimation (LIME) method that estimates the illumination component by finding the maximum RGB values of each pixel and improves the illumination map using the structure of the illumination [20]. Park et al.

proposed a Retinex-based variational optimization-based method using spatially adaptive weight map [38]. Ren et al. proposed a robust low-light enhancement method via low-rank regularized Retinex model (LR3M) by formulating two variational models to estimate illumination and reflectance [39]. Li et al. created an optimization function with a new regularization terms for jointly estimating a piece-wise smoothed illumination and a structure-revealed reflectance [40]. Hao et al. proposed a Retinex-based low-light enhancement method using an efficient semi-decoupled way [21].

In spite of the flexibility using various types of regularization, the optimization-base methods have a common drawback of high computational complexity and long processing time. To solve this problem, various deep learning-based methods were proposed using end-to-end training and a single feedforward inference as summarized in the following subsection.

2.3. Learning-based methods

Ren et al. established a hybrid network that simultaneously estimates global content and salient structures [41]. Ignatov et al. proposed an end-to-end residual network to learn a translation function that improves both color rendition and image sharpness [42]. Park et al. proposed a dual autoencoder network model consisting of stacked and convolutional autoencoders [43]. Park's method performed both low-light enhancement and noise reduction at the same time. Jiang et al. proposed a generative adversarial network (GAN)-based unsupervised model for low-light image enhancement using unpaired low-light and normal-illumination data [44]. Li et al. proposed a trainable convolutional neural network (CNNs) called *Lightennet* to enhance weakly illuminated image [45]. Lore et al. presented a deep autoencoder that simultaneously performs contrast enhancement and denoising [46]. Lv et al. propose MBLLEN proposed a multi-branch low-light image enhancement network (MBLLEN) to extract various features of low-light images with complex contents [14]. Shen et al. proposed a multi-scale Retinex-based CNNs to learn end-to-end mapping between a low-light image and its corresponding enhanced version [47]. Wei et al. proposed an end-to-end deep CNN called Retinex-Net based on the Retinex model [15]. Wei's method consists of decomposition network, enhancement network and denoising operation. The decomposition network decomposes an input low-light image into reflectance and illumination componetns, and the enhancement network brightens up illumination and reduces noise in reflectance. Zero-reference deep curve estimation (Zero-DCE) method showed a lightweight deep neural network for estimating the lighting enhancement curve [48]. The network used non-referenced loss functions to train without any reference images.

3. Methodology

The proposed low-light enhancement framework is combined optimization-based decomposition and learning-based enhancement network (CODEN) as shown in Fig. 2.

The proposed algorithm goes through a two-step process to enhance low-light image. In the decomposition step, the low-light input image, f_{in}^{low}, is decomposed into illumination, f_L^{low}, and reflectance, f_R^{low}, components based on the Retinex model. The enhancement step first creates a set of three derived inputs including f_L^{low}, the gamma corrected illumination, f_L^{GC}, and the histogram equalized illumination, f_L^{HE}. The illumination enhancement net (IEN) then takes as input the set of three derived inputs to enhance the brightness and contrast of the illumination. The enhanced illumination, f_L^E, and f_R^{low} are combined by pixelwise multiplication to produce the enhanced output image, f_{out}^E.

3.1. Decomposition

The decomposition process shown in Fig. 2 combines the Retinex model-based optimization and learning-based decomposition net (DCN) as shown in Fig. 3.
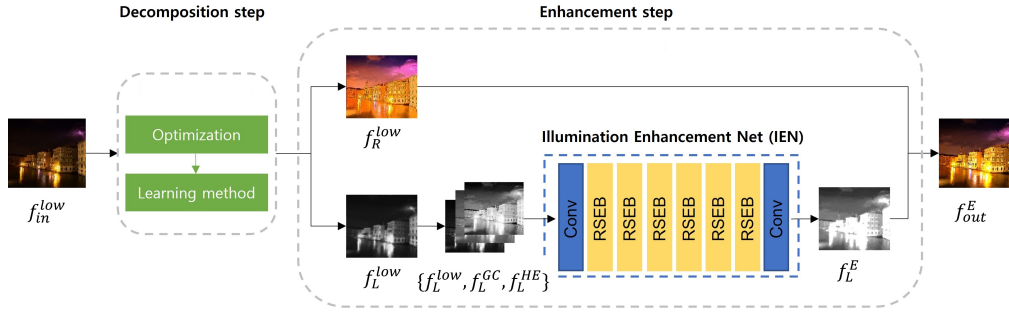
Fig. 2. The combined optimization-based decomposition and enhancement network (CODEN).
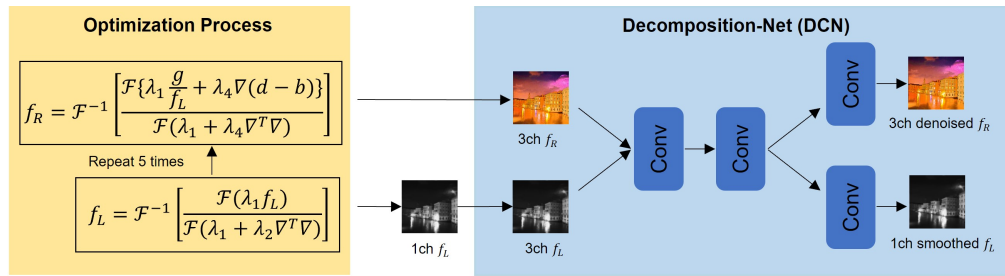

Fig. 3. The decomposition step of the proposed framework. The number of channels are provided for each f_L and f_R.

In the decomposition process, gradients of the illumination are minimized via edge-preserved smoothing, and the reflectance is denoised while preserving detail information. The optimization process enables flexible adjustment of parameters and regularization term [17]. On the other hand, DCN accelerates the entire decomposition process using parallel computation of convolution layers [17]. As a result, we can take advantage of both flexibility by the optimization process and fast processing speed by parallel processing of the convolutional layers.

In the optimization process in Fig. 3, the reflectance and illumination are estimated by minimizing the combined ℓ_1 and ℓ_2 objective function as:

$$\arg\min_{f_L, f_R} \lambda_1 \|f_L f_R - g\|_2^2 + \lambda_2 \|\nabla f_L\|_2^2 + \lambda_3 \|\nabla f_R\|_1, \tag{1}$$

where f_L and f_R respectively represent illumination and reflectance, ∇ is the gradient opertion, and λ_1, λ_2 and λ_3 are regularization parameters for corresponding terms. The first term in Eq. (1) is a data fidelity term for the input image, and minimization of $\|\nabla f_L\|_2^2$ makes the illumination smooth in the ℓ_2 sense. Minimization of $\|\nabla f_R\|_1$ removes noise while preserving edges. Since the ℓ_1-norm is not differentiable, minimization of Eq. (1) is transformed to the differentiable problem using the splitting method [49] as

$$\arg\min_{f_L, f_R} \lambda_1 \|f_L f_R - g\|_2^2 + \lambda_2 \|\nabla f_L\|_2^2 + \lambda_3 \|d\|_1 + \lambda_4 \|d - \nabla f_R - b\|_2^2, \tag{2}$$

where λ_3 and λ_4 represent bregman penalization parameters. The auxiliary variable d and the Bregman variable b are computed as [49]

$$d = \text{shrink}(\nabla f_R + b, \lambda_3/\lambda_5), b = b + \nabla f_R - d, \tag{3}$$

where the shrink operation is defined as

$$\text{shrink}(x, \alpha) = \frac{x}{|x|} \max(x - \alpha, 0), \tag{4}$$

The closed form solutions of (2) can be obtained using the Fourier transfrom as

$$f_L = F^{-1}\left[\frac{F\{\lambda_1 f_L\}}{\lambda_1 + \lambda_2 \nabla^T \nabla}\right], \text{ and } f_R = F^{-1}\left[\frac{F\left\{\lambda_1 \frac{g}{f_L} + \lambda_4 \nabla^T (d - b)\right\}}{F\left(\lambda_1 + \lambda_4 \nabla^T \nabla\right)}\right], \tag{5}$$

where f_L is the smoothed illumination, and f_R is the denoised reflectance with preserved edges. More specifically, f_R is obtained using the pre-estimated f_L. The corresponding optimization process repeats five times until convergence.

The decomposition net (DCN) takes two inputs: i) the reflectance estimated by the optimization method and ii) a set of three derived illumination inputs including, the estimated, histogram equalized, and gamma corrected illuminations. For fast computation and acceptable performance, the DCN has three convolutional layers to generate the denoised reflectance with preserved details and the smoothed illumination with a wide dynamic range and high contrast. To train the DCN, we used edge-preserving smoothing datasets [50] as shown in Fig. 4. Since the reflectance has three channels in the DCN input, we combined three illumination channels estimated by the optimization process. By making the reflectance and illumination inputs have the same number of channels, we can share the first conv layer of the DCN. For learning the illumination, the grayscale of ground truth (GT) was used. On the other hand for learning the reflectance, an additive white Gaussian noise of range $[0, 50]$ was added to the input data.

Fig. 4. Pair of input (top row) and edge-preserving smoothing images (bottom row) in the dataset [50].

The loss function to train the DCN consists of two terms: illumination estimation loss and reflectance estimation loss as

$$\mathcal{L}_{DCN} = \mathcal{L}_{f_L} + \mathcal{L}_{f_R}, \tag{6}$$

where

$$\mathcal{L}_{f_L} = \frac{1}{N}\sum_i^N \left\| f_L - f_{GT}^{gr} \right\|^2, \text{ and } \mathcal{L}_{f_R} = \frac{1}{N}\sum_i^N \left\| f_R - f_{in} \right\|^2, \tag{7}$$

where f_L, f_R, f_{GT}^{gr}, and f_{in} respectively represent the estimated illumination, estimated reflectance, grayscale GT, and input images of the datasets in [50]. Minimization of $\mathcal{L}_{f_L}$ removes texture components from the illumination component and learns GT images from the edge-preserving smoothing dataset [50], whereas minimization of $\mathcal{L}_{f_R}$ removes noise in the reflectance component preserving edge details using the input images in the dataset.

In order to prove the efficiency of our decomposition method, an ablations study was performed on both learning and optimization steps as shown in Fig. 5. For the illumination, the learning step failed in complete smoothing in some textured regions. For the illumination, the learning step failed in complete smoothing in some textured regions, which results in incompletely estimating the details in the reflectance. On the other hand, for the reflectance, the learning step successfully removed noise as shown in the cropped sky region while it could not preserve the details in the cropped cliff region. The optimization step succeeded in globally smoothing the illumination, but lost some important structures, which might result in halo artifacts in the final output image. For the reflectance, the optimization step could not completely remove the noise. On the other hand, the combined optimization and learning steps produced well-smoothed illumination while preserving structures, and well removed noise in the reflectance while preserving edges.
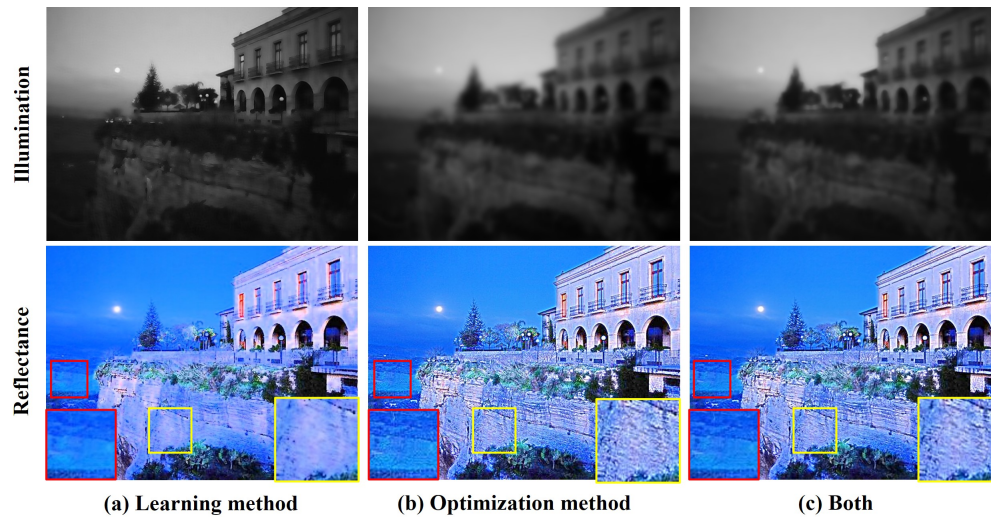

(a) Learning method   **(b) Optimization method**   **(c) Both**

Fig. 5. Comparison of decomposed illumination and reflectance according to ablation study of learning and optimization steps.

3.2. Enhancement

In the enhancement step, the brightness and contrast of the illumination, which is one of the two outputs from the decomposition step, are enhanced. Since there is a limit to enhancing both brightness and contrast with the decomposed illumination alone, gamma correction (GC) and histogram equalization (HE) illuminations are additionally combined to create the set of derived inputs as shown in Fig. 6. The GC input helps improving and adjusting the brightness of the result image, and the HE input helps improving the contrast of the result image.

Results of gamma-corrected images with four different gamma values are shown in Fig. 7. By changing the value of γ in $f_L^{GC} = (f_L^{low})^\gamma$, the gamma value of GC is adjusted. As γ decreases, the brightness of the result image increases.

The illumination enhancement net (IEN) takes as input the set of derived illumination inputs, and produces the resulting illumination with enhanced brightness and contrast. As shown in Fig. 2, the IEN consists of two convolutional layers and six RSEBs. Although Gu *et al.* used 20 RSEBs to improve the performance of remote sensing super-resolution in their original work, we used only six RSEBs to speep up learning without gradient vanishing. We experimentally observed that more than six RSEBs did not further improve the enhancement performance. An RSEB takes local residual learning, which improves learning speed and prevents gradient vanishing as shown in Fig. 8(a). The squeeze and excitation block (SEB) has squeezing and
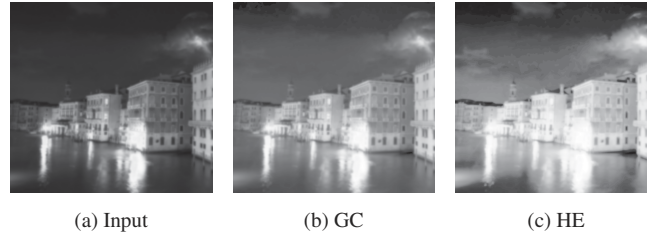
(a) Input   (b) GC   (c) HE

Fig. 6. Three derived inputs: (a) the decomposed illumination, (b) the gamma-corrected (GC) illumination, and (c) the histogram-equalized (HE) illumination.
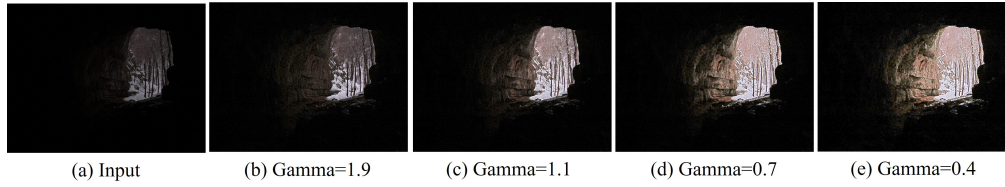

(a) Input   (b) Gamma=1.9   (c) Gamma=1.1   (d) Gamma=0.7   (e) Gamma=0.4

Fig. 7. Comparison of brightness of result image according to gamma value of GC of derived inputs.

(a) Residual Squeeze and Excitation Block (RSEB)   (b) Squeeze and Excitation Block (SEB)
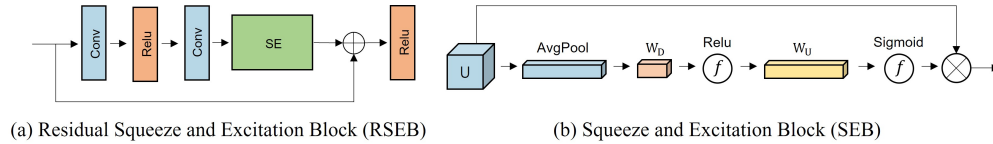
Fig. 8. Residual squeeze and excitation block (RSEB) and squeeze and excitation block (SEB) included in Enhance-Net.

excitation operations [51] as shown in Fig. 8(b). In the squeezing operation, a channel descriptor is created for each channel through average pooling. In the excitation operation, the number of dimensions of the feature map is reduced by the reduction ratio and the importance of the channel is readjusted. After that, the dimension of the feature map is restored to its original size. The SEB can help the proposed network to learn more in the process of compressing and expanding the scalar representing the feature map.

To train the IEN, we used the low-light (LOL) datasets as shown in Fig. 9 [20]. The LOL dataset consist of 500 pairs of low- and normal-light images taken from real scenes including 485 train data and 15 evaluation data. We added 485 synthetic data by applying gamma correction to the normal-light image of train data of LOL datasets to increase the number of train data. A total of 970 paired data were used for training by combining real data and synthetic data.

The loss function for the IEN uses mean squared error (MSE) as

$$\mathcal{L}_{IEN} = \frac{1}{N} \sum_{i}^{N} \left\| f_L^E - f_{GT}^{LOL} \right\|^2, \tag{8}$$

where f_L^E represents the enhanced illumination, that is the output of the IEN, and f_{GT}^{LOL} the normal-light image of the LOL dataset.

To train the IEN, we made derived inputs including gamma-corrected and histogram equalized low-light images, and made the grayscale of normal-light images for the IEN output. Since the IEN enhances only brightness and contrast, it does not perform the decomposition step for the LOL dataset.
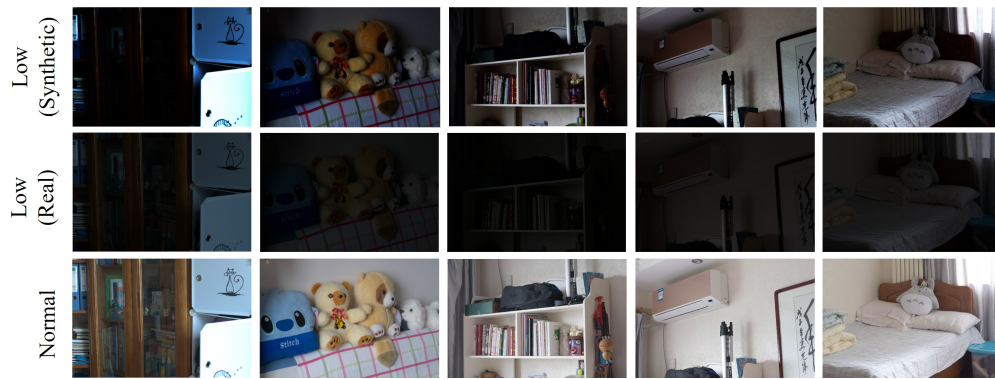
Fig. 9. Normal / Low(real) / Low(synthetic)-light pairs of the LOL dataset that is available publicly [15].

4. Experiment results

The proposed method is compared with the existing method from a qualitative and quantitative point of view through extensive experiments. For comparison, we use the published code for the existing methods. Overall, we conducted the experiment from three perspectives. 1) We compare the proposed method with existing methods including optimization methods, deep learning methods, contrast enhancement methods, etc. in the low-light enhancement task. 2) We prove the superiority of the proposed method through qualitative and quantitative evaluation. 3) We explain the necessity of each element of the proposed method through an ablation study.

The optimization method and learning method used in this paper were implemented in Python 3.8. We also used pytorch open source machine learning framework for deep learning model inference. All the algorithms were run on Nvidia GeForce RTX 3090 GPU and Intel Core i7-8700K CPU equipped with 88 GB of RAM.

For training Decom-Net and Enhance-Net, adam optimizer is used for back-propagation, learning rate is set to be $1e^{-4}$, batch size is set to be 4 and patch-size to be 224 x 224. The (6), (8) loss functions are used to train Decom-Net and Enhance-Net penalizing the network for creating output that similar to target. The learning-based methods are trained on Nvidia GeForce RTX 3090 GPU. The training code ran for about 3 hours to train Decom-Net and about 6 hours to train Enhance-Net. In decomposition step, when model-based optimization is used, λ_1, λ_2, λ_3 and λ_4 are set to 100, 100, 1 and 1 respectively.

4.1. Ablation study

In this paper, we propose Enhance-Net to improve the brightness and contrast of decomposed illumination based on the Retinex model. Enhance-Net includes derived inputs and RSEB. Figure 10 shows the results of performing an ablation study for two features. In the case of removing the derived inputs, it was replaced with three channels of same decomposed illumination, and in the case of removing the RSEB, that was replaced with a plain CNNs with the same number of layers. In the case of (b), when all are removed, the result shows that the brightness is not significantly improved. In the case of (c), when only the derived inputs were used, the brightness was greatly improved, but the contrast was not preserved. In the case of (d), when only RSEB is used, the contrast is excessively improved, showing the result that the detail of the bird's wing is saturated. In the case of (e), when both are used, both the brightness and contrast are properly improved, showing a visually excellent result. The blind/referenceless image spatial quality evaluator (BRISQUE) [52] metric was used to quantitatively evaluate the result of each

case. The BRISQUE is a method to numerically measure the degree of distortion of natural images through mean subtracted contrast normalized (MSCN) processing. Accordingly, it can be said that the lower the BRISQUE score, the closer to a natural image without distortion. The case of (e) using both RSEB and derived inputs shows the lowest BRISQUE score.
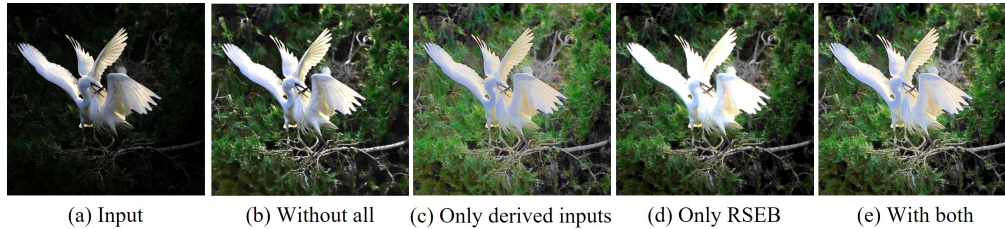

(a) Input   (b) Without all   (c) Only derived inputs   (d) Only RSEB   (e) With both

Fig. 10. Comparison of brightness and contrast of the resulting image according to the ablation study of the derived inputs and RSEB: (b) Without all (BRISQUE:24.07), (c) Only derived inputs (BRISQUE:15.68), (d) Only RSEB (BRISQUE:27.05), (e) With both (BRISQUE:10.98)

4.2. Visual and subjective comparisons

For the qualitative evaluation of the proposed method, a comparison of the results of several different methods is shown in Fig. 11. Figure 11 presents a comparison of three natural low-light images. When viewed as a whole or in detail, our method achieves visually pleasing results by showing enhanced results in both brightness and contrast with fewer artifacts. As shown in the part cropped by the red square, our method improves the brightness for the part where the dark and bright regions of the input image are together, while preserving the contrast well. On the other hand, CVC and LDR methods do not improve brightness significantly because they only improve contrast [12,13]. Since MBLLEN and Retinex-Net methods are brightness enhancement methods by deep learning networks, they show a result that does not preserve contrast by significantly improving the brightness of dark areas [14,15]. In addition, the Retinex-Net method and Park's method have limitations in smoothing the texture during decomposition based on the Retinex model if you look at the cropped image, you can see that the detail of the texture part of the result image is lost [15,38]. If the illumination is further smoothed for preserving detail in the texture part of the result image, a halo effect can be occurred in the structure part of the result image. However, proposed method smoothes the edges while preserving the edges of the illumination when decomposed based on Retinex model, thereby preserves the detail of the texture part of the result image and minimizes the halo effect of the structure part.

4.3. Quantitative comparisons

We demonstrate the objective performance of ours using quantitative comparison of the proposed method with 11 other methods (CVC, LDR, MBLLEN, Zero-DCE, LIME, Retinex-Net, Park et al., and Hao et al.) A fusion-based enhancing method (MF), robust Retinex model (RRM), LECARM) using two kind of metrics, no reference image quality assessment (IQA) and full reference IQA [12–15,20,21,37,40,43,48,53]. No-reference IQA requires only one image for quality measurement, and full reference IQA requires a pair of images for quality measurement and original images.

Table 1 shows values measured by natural image quality evaluator (NIQE) and perception-based image quality evaluator (PIQE) metrics, that are non-reference measurement IQA [54,55]. The NIQE is predicted as the gap between a multivariate Gaussian (MVG) fit of the NSS features derived from the test image, and a MVG model of the quality aware features derived from clean images without distortion. The PIQE is a measurement method that indicates the degree of
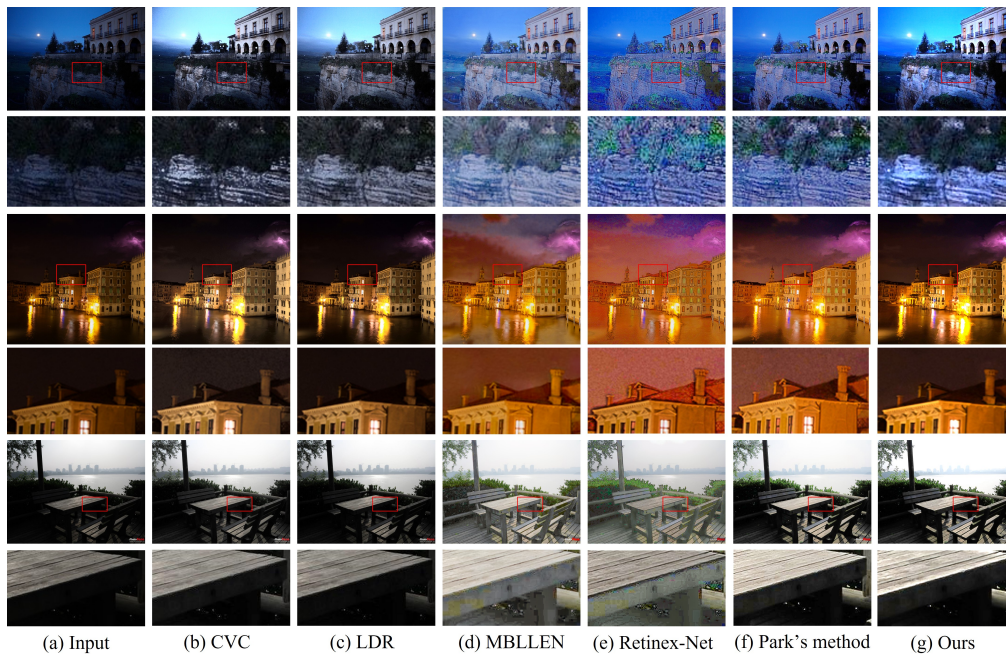
| (a) Input | (b) CVC | (c) LDR | (d) MBLLEN | (e) Retinex-Net | (f) Park's method | (g) Ours |

Fig. 11. Comparison with other methods for images with dark area.

distortion of a test image by using the human visual system. Therefore, the lower the NIQE and PIQE scores, the better the image quality. A total of 40 data from DICM, Fusion, LIME, and MEF datasets were used for measurement. It can be seen that ours has the lowest values for both NIQE and PIQE.

Table 1. Average NIQE, PIQE value for randomly selected 40 images of DICM, Fusion, LIME, MEF datasets [13,20,55,57,58].

Metrics	CVC	LDR	MBLLEN	Zero-DCE	LIME	Retinex-Net
NIQE	3.331	3.215	3.008	3.390	3.283	4.927
PIQE	35.603	33.637	31.35	33.456	44.243	39.838
Metrics	Park et al.	Hao et al.	MF	RRM	LECARM	Ours
NIQE	3.689	3.267	3.139	3.836	3.261	**2.862**
PIQE	46.822	38.396	32.618	41.168	31.903	**28.410**

Table 2 shows the values measured by PSNR and SSIM metric, which are full reference IQA [56]. PSNR is a measurement method representing the value of noise with respect to the maximum value that a signal can have. SSIM is a method of measuring the degree of structural distortion by comparing a test image with an original image. The higher PSNR and SSIM values, the better image quality. For measurement data, 15 test images of LOL datasets were used. Also, it can be seen that ours has the highest values for both PSNR and SSIM.

Table 2. Average PSNR, SSIM value for 15 test images of LOL datasets.

Metrics	CVC	LDR	MBLLEN	Zero-DCE	LIME	Retinex-Net
PSNR	10.48	14.443	12.863	13.547	14.521	15.145
SSIM	0.23	0.433	0.431	0.449	0.325	0.325
Metrics	Park et al.	Hao et al.	MF	RRM	LECARM	Ours
PSNR	13.028	12.046	15.236	12.514	12.863	**15.394**
SSIM	0.395	0.528	0.389	0.567	0.431	**0.577**

5. Conclusion

To enhance low-light images by optimally adjusting both brightness and contrast, we combined optimization-based decomposition and learning-based enhancement network. The proposed method is divided into decomposition and enhancement steps. In the decomposition step, an input image is decomposed into illumination and reflectance using both optimization and learning methods. In the enhancement step, the brightness and contrast of the decomposed illumination are simultaneously enhanced by the illumination enhancement net (IEN). The proposed method improves low-light images in an end-to-end manner. The several experiments showed the effectiveness of each component of the proposed framework. The proposed DCN and IEN used paired low/normal-light datasets for training with additional synthetic data. In the future, we will try to design an unsupervised learning method using unpaired images.

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

References

1. N. Choi, J. Jang, and J. Paik, "Illuminant-invariant stereo matching using cost volume and confidence-based disparity refinement," J. Opt. Soc. Am. A **36**(10), 1768–1776 (2019).
2. J. Ju, D. Kim, B. Ku, D. K. Han, and H. Ko, "Online multi-object tracking with efficient track drift and fragmentation handling," J. Opt. Soc. Am. A **34**(2), 280–293 (2017).
3. X. Wang, S. Shen, C. Ning, Y. Zhang, and G. Lv, "Robust object tracking based on local discriminative sparse representation," J. Opt. Soc. Am. A **34**(4), 533–544 (2017).
4. W. Cho, J. Jang, A. Koschan, M. A. Abidi, and J. Paik, "Hyperspectral face recognition using improved inter-channel alignment based on qualitative prediction models," Opt. Express **24**(24), 27637–27662 (2016).
5. Y.-S. Chen, D. Meng, W.-Z. Ma, W. Chen, P.-P. Zhuang, W. Chen, Z.-C. Fan, C. Dou, Y. Gu, and J. Liu, "Fingerprint detection in the mid-infrared region based on guided-mode resonance and phonon-polariton coupling of analyte," Opt. Express **29**(23), 37234–37244 (2021).
6. X.-F. Liu, X.-R. Yao, R.-M. Lan, C. Wang, and G.-J. Zhai, "Edge detection based on gradient ghost imaging," Opt. Express **23**(26), 33802–33811 (2015).
7. H. Lim, S. Yu, K. Park, D. Seo, and J. Paik, "Texture-aware deblurring for remote sensing images using ℓ_0-based deblurring and ℓ_2-based fusion," IEEE J. Sel. Top. Appl. Earth Observations Remote Sensing **13**, 3094–3108 (2020).
8. J. Shin, H. Park, and J. Paik, "Region-based dehazing via dual-supervised triple-convolutional network," IEEE Trans. Multimedia **24**, 245–260 (2021).
9. R. Khan, Y. Yang, Q. Liu, J. Shen, and B. Li, "Deep image enhancement for ill light imaging," J. Opt. Soc. Am. A **38**(6), 827–839 (2021).
10. M. Lecca, "Generalized equation for real-world image enhancement by milano retinex family," J. Opt. Soc. Am. A **37**(5), 849–858 (2020).
11. W. Wang, C. Zhang, and M. K. Ng, "Variational model for simultaneously image denoising and contrast enhancement," Opt. Express **28**(13), 18751–18777 (2020).
12. T. Celik and T. Tjahjadi, "Contextual and variational contrast enhancement," IEEE Trans. on Image Process. **20**(12), 3431–3441 (2011).
13. C. Lee, C. Lee, and C.-S. Kim, "Contrast enhancement based on layered difference representation of 2d histograms," IEEE Trans. on Image Process. **22**(12), 5372–5384 (2013).
14. F. Lv, F. Lu, J. Wu, and C. Lim, "Mbllen: Low-light image/video enhancement using cnns," in *BMVC*, vol. 220 (2018).

15. C. Wei, W. Wang, W. Yang, and J. Liu, "Deep retinex decomposition for low-light enhancement," arXiv preprint arXiv:1808.04560 (2018).

16. E. H. Land, "The retinex theory of color vision," Sci. Am. **237**(6), 108–128 (1977).

17. K. Zhang, W. Zuo, S. Gu, and L. Zhang, "Learning deep cnn denoiser prior for image restoration," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, (2017), pp. 3929–3938.

18. W. Ren, L. Ma, J. Zhang, J. Pan, X. Cao, W. Liu, and M.-H. Yang, "Gated fusion network for single image dehazing," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, (2018).

19. J. Gu, X. Sun, Y. Zhang, K. Fu, and L. Wang, "Deep residual squeeze and excitation network for remote sensing image super-resolution," Remote Sens. **11**(15), 1817 (2019).

20. X. Guo, Y. Li, and H. Ling, "Lime: Low-light image enhancement via illumination map estimation," IEEE Trans. on Image Process. **26**(2), 982–993 (2016).

21. S. Hao, X. Han, Y. Guo, X. Xu, and M. Wang, "Low-light image enhancement with semi-decoupled decomposition," IEEE transactions on multimedia **22**(12), 3025–3038 (2020).

22. M. Abdullah-Al-Wadud, M. H. Kabir, M. A. A. Dewan, and O. Chae, "A dynamic histogram equalization for image contrast enhancement," IEEE Trans. Consumer Electron. **53**(2), 593–600 (2007).

23. S. Wang, W. Cho, J. Jang, M. A. Abidi, and J. Paik, "Contrast-dependent saturation adjustment for outdoor image enhancement," J. Opt. Soc. Am. A **34**(1), 7–17 (2017).

24. D. J. Jobson, Z.-u. Rahman, and G. A. Woodell, "Properties and performance of a center/surround retinex," IEEE Trans. on Image Process. **6**(3), 451–462 (1997).

25. 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 Trans. on Image Process. **6**(7), 965–976 (1997).

26. C.-H. Lee, J.-L. Shih, C.-C. Lien, and C.-C. Han, "Adaptive multiscale retinex for image contrast enhancement," in *2013 International Conference on Signal-Image Technology & Internet-Based Systems*, (IEEE, 2013), pp. 43–50.

27. S. Wang, J. Zheng, H.-M. Hu, and B. Li, "Naturalness preserved enhancement algorithm for non-uniform illumination images," IEEE Trans. on Image Process. **22**(9), 3538–3548 (2013).

28. L. Wang, L. Xiao, H. Liu, and Z. Wei, "Variational bayesian method for retinex," IEEE Trans. on Image Process. **23**(8), 3381–3396 (2014).

29. G. Deng, "A generalized unsharp masking algorithm," IEEE Trans. on Image Process. **20**(5), 1249–1261 (2010).

30. L. Li, R. Wang, W. Wang, and W. Gao, "A low-light image enhancement method for both denoising and contrast enlarging," in *2015 IEEE International Conference on Image Processing (ICIP)*, (IEEE, 2015), pp. 3730–3734.

31. L. Yuan and J. Sun, "Automatic exposure correction of consumer photographs," in *European Conference on Computer Vision*, (Springer, 2012), pp. 771–785.

32. Q. Shan, J. Jia, and M. S. Brown, "Globally optimized linear windowed tone mapping," IEEE Trans. Visual. Comput. Graphics **16**(4), 663–675 (2009).

33. B. Cai, X. Xu, K. Guo, K. Jia, B. Hu, and D. Tao, "A joint intrinsic-extrinsic prior model for retinex," in *Proceedings of the IEEE international conference on computer vision*, (2017), pp. 4000–4009.

34. X. Fu, Y. Liao, D. Zeng, Y. Huang, X.-P. Zhang, and X. Ding, "A probabilistic method for image enhancement with simultaneous illumination and reflectance estimation," IEEE Trans. on Image Process. **24**(12), 4965–4977 (2015).

35. X. Fu, D. Zeng, Y. Huang, P. Zhang, and X. Ding, "A weighted variational model for simultaneous reflectance and illumination estimation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, (2016), pp. 2782–2790.

36. X. Ren, M. Li, W.-H. Cheng, and J. Liu, "Joint enhancement and denoising method via sequential decomposition," in *2018 IEEE international symposium on circuits and systems (ISCAS)*, (IEEE, 2018), pp. 1–5.

37. Y. Ren, Z. Ying, T. H. Li, and G. Li, "Lecarm: Low-light image enhancement using the camera response model," IEEE Trans. Circuits Syst. Video Technol. **29**(4), 968–981 (2018).

38. S. Park, S. Yu, B. Moon, S. Ko, and J. Paik, "Low-light image enhancement using variational optimization-based retinex model," IEEE Trans. Consumer Electron. **63**(2), 178–184 (2017).

39. X. Ren, W. Yang, W.-H. Cheng, and J. Liu, "Lr3m: Robust low-light enhancement via low-rank regularized retinex model," IEEE Trans. on Image Process. **29**, 5862–5876 (2020).

40. M. Li, J. Liu, W. Yang, X. Sun, and Z. Guo, "Structure-revealing low-light image enhancement via robust retinex model," IEEE Trans. on Image Process. **27**(6), 2828–2841 (2018).

41. W. Ren, S. Liu, L. Ma, Q. Xu, X. Xu, X. Cao, J. Du, and M.-H. Yang, "Low-light image enhancement via a deep hybrid network," IEEE Trans. on Image Process. **28**(9), 4364–4375 (2019).

42. A. Ignatov, N. Kobyshev, R. Timofte, K. Vanhoey, and L. Van Gool, "Dslr-quality photos on mobile devices with deep convolutional networks," in *Proceedings of the IEEE International Conference on Computer Vision*, (2017), pp. 3277–3285.

43. S. Park, S. Yu, M. Kim, K. Park, and J. Paik, "Dual autoencoder network for retinex-based low-light image enhancement," IEEE Access **6**, 22084–22093 (2018).

44. Y. Jiang, X. Gong, D. Liu, Y. Cheng, C. Fang, X. Shen, J. Yang, P. Zhou, and Z. Wang, "Enlightengan: Deep light enhancement without paired supervision," IEEE Trans. on Image Process. **30**, 2340–2349 (2021).

45. C. Li, J. Guo, F. Porikli, and Y. Pang, "Lightennet: A convolutional neural network for weakly illuminated image enhancement," Pattern recognition letters **104**, 15–22 (2018).

46. K. G. Lore, A. Akintayo, and S. Sarkar, "Llnet: A deep autoencoder approach to natural low-light image enhancement," Pattern Recognition **61**, 650–662 (2017).

47. L. Shen, Z. Yue, F. Feng, Q. Chen, S. Liu, and J. Ma, "Msr-net: Low-light image enhancement using deep convolutional network," arXiv preprint arXiv:1711.02488 (2017).

48. C. Guo, C. Li, J. Guo, C. C. Loy, J. Hou, S. Kwong, and R. Cong, "Zero-reference deep curve estimation for low-light image enhancement, in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, (2020), pp. 1780–1789.

49. T. Goldstein and S. Osher, "The split bregman method for ℓ_1-regularized problems," SIAM journal on imaging sciences **2**(2), 323–343 (2009).

50. F. Zhu, Z. Liang, X. Jia, L. Zhang, and Y. Yu, "A benchmark for edge-preserving image smoothing," IEEE Trans. on Image Process. **28**(7), 3556–3570 (2019).

51. J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, (2018).

52. A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," IEEE Trans. on Image Process. **21**(12), 4695–4708 (2012).

53. X. Fu, D. Zeng, Y. Huang, Y. Liao, X. Ding, and J. Paisley, "A fusion-based enhancing method for weakly illuminated images," Signal Processing **129**, 82–96 (2016).

54. A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a completely blind image quality analyzer," IEEE Signal Process. Lett. **20**(3), 209–212 (2013).

55. N. Venkatanath, D. Praneeth, Bh Maruthi Chandrasekhar, S. Channappayya Sumohana, and Swarup S. Medasani, "Blind image quality evaluation using perception based features," in *2015 Twenty First National Conference on Communications (NCC)*, (2015), pp. 1–6.

56. Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: from error visibility to structural similarity," IEEE Trans. on Image Process. **13**(4), 600–612 (2004).

57. K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising with block-matching and 3d filtering," in *Image Processing: Algorithms and Systems, Neural Networks, and Machine Learning*, vol. 6064 (International Society for Optics and Photonics, 2006), p. 606414.

58. K. Ma, K. Zeng, and Z. Wang, "Perceptual quality assessment for multi-exposure image fusion," IEEE Trans. on Image Process. **24**(11), 3345–3356 (2015).