Luminance-aware Color Transform for Multiple Exposure Correction

Jong-Hyeon Baek† † DaeHyun Kim† † Su-Min Choi2 Hyo-jun Lee1 Hanul Kim3 Yeong Jun Koh1∗
1Chungnam National University 242dot Inc.
3Seoul National University of Science and Technology

whdgusd1978@gmail.com, seven776484@gmail, sumin.choi@42dot.ai,
gwyns6287@gmail.com, hukim@seoultech.ac.kr, yjkoh@cnu.ac.kr

Abstract

Images captured with irregular exposures inevitably present unsatisfactory visual effects, such as distorted hue and color tone. However, most recent studies mainly focus on underexposure correction, which limits their applicability to real-world scenarios where exposure levels vary. Furthermore, some works to tackle multiple exposure rely on the encoder-decoder architecture, resulting in losses of details in input images during down-sampling and up-sampling processes. With this regard, a novel correction algorithm for multiple exposure, called luminance-aware color transform (LACT), is proposed in this study. First, we reason the relative exposure condition between images to obtain luminance features based on a luminance comparison module. Next, we encode the set of transformation functions from the luminance features, which enable complex color transformations for both overexposure and underexposure images. Finally, we project the transformed representation onto RGB color space to produce exposure correction results. Extensive experiments demonstrate that the proposed LACT yields new state-of-the-arts on two multiple exposure datasets. Code is available at https://github.com/whdgusd148/LACT.

1. Introduction

Exposure is a crucial component in image formation. So, images captured with improper exposures usually have unsatisfactory visual effects, such as distorted hue and brightness. Figure 1 shows examples of images captured with overexposure and underexposure, which can cause overly bright or darkened regions, leading to significant degradation of image quality. Many methods [3, 21, 1, 22, 8, 23, 41, 20] have been proposed to correct exposure-related issues. However, most of them have focused on correcting either overexposure or underexposure, limiting their applicability to real-world scenarios with various exposure levels. Some studies [2, 14, 35, 15] have attempted to address both overexposure and underexposure cases, leveraging the U-Net [31] like architecture. While these methods have shown improved performance, they still suffer from limitations such as the down-sampling and up-sampling processes in their approaches, resulting in the loss of image details.

The transformation function-based approach has gained much attention in image enhancement [10, 20, 19, 39]. This approach estimates transformation functions for color channels and performs color transformation based on the predicted functions. These methods are free from the down-sampling and up-sampling processes, thereby preserving the image details. However, they are still vulnerable to multiple exposures, since they do not exploit exposure information explicitly. Also, as shown in Figure 1, a single transformation function may not be sufficient to address both overexposed or underexposed images, particularly those containing saturated regions.

In this work, we present a novel approach, named luminance-aware color transform (LACT), to tackle the problem of multiple exposure correction. The proposed method leverages two key insights - the critical role of luminance for correcting images with multiple exposure levels, and the need to overcome the limitation of single intensity transformation function. To this end, the proposed method extracts luminance features, which encode exposure levels in input images, by exploiting the luminance order relationship between two randomly sampled input images. The luminance features are used in the luminance-aware color transform to estimate the sets of transformation functions from the local structure, enabling complex color transform for both overexposure and underexposure images. Then, the proposed method generates a multi-channel representation for exposure correction using the set of transformation functions. Finally, a post-processing module projects this multi-channel representation onto the RGB color space to produce correction results. Extensive experiments on ME [2] and

The main contributions of our work are three folds:

- We propose the luminance comparison module that leverages the luminance order relationship to extract the effective luminance feature.
- We propose the luminance-aware color transform based on the luminance feature to obtain reliable correction results for both underexposure and overexposure images.

2. Related Work

2.1. Exposure correction

Early studies in exposure correction mainly aimed at enhancing the global contrast of an input image. Histogram equalization [28] and its modifications [1, 30, 27] are a well-known technique that modifies an image’s histogram to stretch its limited contrast. Another popular method is the power-law (gamma) transformation [9, 16, 28], which maps input pixel values to output pixel values using predefined transformations. Retinex theory [21]-based methods [11, 18, 25, 36] decompose exposure correction into illumination enhancement and reflectance regularization.

Recently, deep learning-based methods [7, 41, 40, 38] have been proposed for exposure correction. For instance, SID [7] trains a fully-convolutional network to process low-light sensing images. In [38], frequency decomposition is adopted for underexposure correction. DBRN [40] decomposes image features into different bands and recursively recomposes them to enhance underexposed images. STAR [41] introduces a structure-aware lightweight Transformer [34] to extract long-short range contexts of low light images. Another approach to exposure correction [19, 20, 10] enhances irregular colors based on an intensity transformation function. Zero-DCE [10] formulates exposure correction as a task of image-specific curve estimation with deep neural networks. GEN-LEN [20] applies global intensity transformation functions and then locally revises the global correction with spatial filtering. RCTN [19] implements color transformation to enhance low-light images based on the similarity between the input image and its representative colors. However, most of these methods focus on underexposure correction, which limits their applications for various exposure levels.

2.2. Multiple exposure correction

Afifi et al. [2] introduced the ME [2] dataset, which includes over 20,000 images with multiple exposure scenes. This has led to lots of attempts [26, 2, 5, 15, 35] to correct exposed images in multiple exposure levels. Wang et al. [35] proposed the local color distribution embedded module to detect overexposed and underexposed regions based on the guidance of the local color distributions. ENC [14] extents the previous underexposure correction [7, 40] to the multiple exposures one by adding a normalization module that embeds multiple exposure features to the exposure-invariant space. FECNet [15] converts the different exposure images into frequency properties, which provide detailed information about lightness and structure for the exposure correction networks. CMEC [26] applies stacked self-attention layers [34] to model interactions between pixels with a diverse range of exposures. These methods [26, 2, 5, 15, 35], leverage U-Net [31] like architecture and have shown promising performance. However, the down-sampling and up-sampling processes in their approaches may result in the loss of image details. In contrast, the proposed LACT aims to estimate transformation functions that are free from these processes, preserving image details while correcting unnatural exposures.
3. Proposed Method

Let \( X \in \mathbb{R}^{H \times W \times 3} \) be an input RGB image captured with either underexposure or overexposure settings, where \( H \) and \( W \) denote the image height and width, respectively. The aim of multiple exposure correction is to restore the high-quality image \( Y \in \mathbb{R}^{H \times W \times 3} \) by compensating for exposure errors in the input image \( X \). However, due to the distinct nature of underexposure and overexposure, different correction procedures are necessary for each exposure case. Therefore, we present a multiple exposure correction network based on the luminance-aware color transform (LACT), which adaptively enhances the input image by considering its exposure condition explicitly.

Figure 2 illustrates the overall training process of the luminance-aware multiple exposure correction network, which consists of four main components: backbone, luminance comparison module, luminance-aware color transform, and post-processing module. The backbone network encodes the input image into image features. The luminance comparison module takes early features of the backbone and extracts luminance features to exploit the exposure conditions of the input image. The luminance-aware color transform combines the luminance features and the image features to estimate the set of transformation functions for each color channel. Then, it performs multi-channel color transformation based on the transformation functions to correct exposure in the input image. Finally, the post-processing module further enhances the image to yield visually pleasing results.

3.1. Backbone

Table 1 presents the detailed architecture of the backbone, which consists of three convolution blocks and four inverted residual blocks [13]. Each convolution block sequentially includes a convolution operation, a batch normalization [17], and a swish activation [29]. IR-block is an inverted residual block [13]. Given the input image \( X \in \mathbb{R}^{H \times W} \), the backbone first resizes it to have the fixed spatial resolution of \( h \times w \), where \( h \) and \( w \) denote the resized height and width. The backbone then feeds the resized input through a single convolution block to extract an intermediate image feature \( Z \in \mathbb{R}^{H/2 \times W/2 \times \tilde{c}} \), where \( \tilde{c} \) is the feature dimension. Next, the backbone transforms the intermediate features into an image feature \( Z \in \mathbb{R}^{H/4 \times W/4 \times \tilde{c}} \) using four inverted residual blocks (IR-block) and two convolution blocks, where \( c \) is the feature dimension.

3.2. Luminance comparison module

For effective multiple exposure correction, it needs to analyze the exposure levels of input images. However, it is hard to directly identify exposure levels from arbitrary images. In contrast, luminance information, which is easily derived from images, can be an important cue for repre-
representing the various exposure levels of images. Thus, we design the proposed network to extract effective luminance features from input images. Inspired by order learning [24], we compare two images to estimate which image is brighter or darker through the luminance comparison module. The proposed luminance comparison module contains the encoder, which is composed of seven convolutional blocks, to extract the luminance feature \( L \in \mathbb{R}^{16 \times 16 \times c} \).

Figure 2 illustrates the proposed training strategy based on the luminance comparison module to relate the luminance feature to exposure levels. Specifically, two different training images, \( X_i \) and \( X_j \), are randomly sampled and their features \( \tilde{Z}_i \) and \( \tilde{Z}_j \) are fed into the encoder in the luminance comparison module to obtain the luminance features \( L_i \) and \( L_j \), respectively. We then employ an auxiliary network to estimate the order relationship between \( X_i \) and \( X_j \). The auxiliary network takes \( L_i \) and \( L_j \), and sequentially applies a global average pooling, a fully connected layer, and concatenation operations to them. Finally, from the concatenated feature, the auxiliary network predicts the order relationship \( r_{ij} \in [0, 1] \) through a fully connected layer with sigmoid activation, where \( r_{ij} \) indicates which image is brighter or darker. In this work, the ground-truth order relationship \( \hat{r}_{ij} \) is set to 1 when \( X_i \) is brighter than \( X_j \), while \( \hat{r}_{ij} \) is set to 0 when \( X_i \) is darker than \( X_j \). The luminance comparison module is trained to minimize the cross-entropy loss between \( r_{ij} \) and \( \hat{r}_{ij} \). Note that the auxiliary network is used in training only, which requires two input images.

In contrast, only one input image is used during inference, since its luminance feature can be extracted from the encoder in the luminance comparison module.

3.3. Luminance-aware color transform

To obtain reliable correction results for both underexposure and overexposure images, the LACT module includes a luminance-aware attention block to effectively convey the luminance-related information to the image. The luminance-aware attention block explicitly interweaves the luminance feature \( L \) into the image feature \( Z \) using the multi-head cross-attention (MCA) mechanism and then estimates transformation functions based on the standard multi-head self-attention (MSA) [34]. MCA is similar to MSA, but it takes the queries and key-value pairs from different sources. Here, we linearly transform the image feature \( Z \) into the query and the luminance feature \( L \) into the key and value.

Let \( n = \frac{h}{64} \cdot \frac{w}{64} \) be the spatial dimension of the image feature. Given \( Z \) and \( L \), we define the luminance-aware image feature \( F = [f_1, f_2, \cdots, f_n] \in \mathbb{R}^{n \times c} \) as

\[
F = LN(Z^* + MCA(Z^*, L^*))
\]

where \( LN(\cdot) \) refers to a layer normalization [4], and \( Z^* \in \mathbb{R}^{n \times c} \) and \( L^* \in \mathbb{R}^{\frac{h}{64} \times \frac{w}{64} \times c} \) are the reshaped image and luminance features, respectively.

It is noteworthy that we take into consideration the local structure of the input image since \( f_m \) relates to the \( m \)-th position of the image feature. Underexposed or overexposed images may contain locally saturated regions, where most pixel intensities in the specific color channel are distributed around 0 or 255. These regions may require different transformation functions to enhance narrow ranges of input intensities. Thus, it is essential to consider these local structures, which spatially adjust image regions, for more accurate multiple exposure correction.

Next, we estimate the multi-channel color transformation set \( T \in \mathbb{R}^{n \times 256} \) by

\[
T = LN(F + MSA(F))
\]

\[
T = MLP(T).
\]

Here, MLP(\cdot) includes three fully connected layers with GELU non-linearity [12]. The estimated set is decomposed into three groups of channel-wise intensity transformation functions:

\[
T = [T_r, T_g, T_b]
\]

where \( T_r \in \mathbb{R}^{256 \times 256}, T_g \in \mathbb{R}^{256 \times 256}, T_b \in \mathbb{R}^{256 \times 256} \) are the set of transformation functions for red, green, blue channels, respectively. Thus, there are \( \frac{n^2}{3} \) transformation functions for each color channel.

We perform channel-wise intensity transform to correct the exposure of input image \( X \) using \( T \). For red channel, for example, each column vector \( t_r \in [t_0, t_1, \ldots, t_{255}]^T \) in \( T_r \) is regarded as a transformation function, which maps \( k \) in \( X \) to \( t_k \) in the output image. We perform channel-wise intensity transform for all transformation functions in \( T_r \), and then stack the outputs to obtain a multi-channel enhanced feature \( \tilde{Y}_r \in \mathbb{R}^{H \times W \times 256} \) for red channel. After repeating this process for all color channels, the luminance-aware color transform module provides the multi-channel enhanced feature, which is given by

\[
\tilde{Y}_L = [\tilde{Y}_r, \tilde{Y}_g, \tilde{Y}_b].
\]

3.4. Post-processing module

The multi-channel representation \( \tilde{Y} \) is obtained by exploiting various transform functions from the local structure of the input image. However, it should be projected onto the RGB color space to restore a high-quality image. As depicted in Figure 2, the post-processing module is designed at the end of the network, which yields a refined correction feature \( \tilde{Y}_P \) using sequential stages. Finally, it produces a final correction image \( Y \) by a weighted sum of \( \tilde{Y}_L \) and \( \tilde{Y}_P \), which is given by

\[
Y = Conv(\alpha_1 \tilde{Y}_L + \alpha_2 \tilde{Y}_P)
\]
where $\alpha_1$ and $\alpha_2$ are learnable weights, and Conv represents the convolution block. Table 2 specifies the detailed architecture of the post-processing module.

### 3.5. Loss functions

#### Luminance ordering loss

We use the luminance ordering loss $L_1$ to train the luminance comparison module. To define the luminance ordering loss, we first convert each input RGB image $X_i$ into a grayscale image $G_i$:

$$G_i = 0.299X_{i,r} + 0.587X_{i,g} + 0.114X_{i,b}$$  \hspace{1cm} (7)

where $X_{i,r}$, $X_{i,g}$, $X_{i,b}$ denote intensity images for red, green, and blue channels, respectively. Next, we randomly sample two images $X_i$ and $X_j$ and set the ground-truth order relationship $\hat{r}_{ij}$ between $X_i$ and $X_j$ as

$$\hat{r}_{ij} = \begin{cases} 1 & \text{if } G_i - G_j \geq 0 \\ 0 & \text{otherwise} \end{cases}$$  \hspace{1cm} (8)

where $\bar{G}_i$ is an average of gray values in $G_i$. Then, we define the luminance ordering loss $L_1$:

$$L_1 = \hat{r}_{ij} \cdot \log(r_{ij}) + (1 - \hat{r}_{ij}) \cdot \log(1 - r_{ij})$$  \hspace{1cm} (9)

as the binary cross-entropy between the predicted order relationship $r_{ij}$ and the ground-truth $\hat{r}_{ij}$.

#### Exposure correction loss

We design the exposure correction loss $L_e$ to penalize discrepancies between the predicted exposure correction image $Y$ and the ground-truth image $\hat{Y}$ in various aspects. $L_e$ is defined as a weighted sum of three loss terms: reconstruction loss $L_c$, frequency loss $L_f$, and perceptual loss $L_p$:

$$L_e = L_c + \lambda_f L_f + \lambda_p L_p.$$  \hspace{1cm} (10)

Here, we set the balancing parameters $\lambda_f$ and $\lambda_p$ to 0.5 and 0.05 to scale the relative importance of each term. The reconstruction loss $L_c$ measures the $l_1$ distance between the predicted image $Y$ and its ground-truth $\hat{Y}$ in the color space:

$$L_c = \|Y - \hat{Y}\|_1.$$  \hspace{1cm} (11)

The frequency loss $L_f$, on the other hand, quantifies the difference between the predicted and ground-truth ones in the frequency space using the 2D fast Fourier transform $\mathcal{F}(\cdot)$:

$$L_f = \|\mathcal{F}(Y) - \mathcal{F}(\hat{Y})\|_1.$$  \hspace{1cm} (12)

$L_f$ is crucial for preserving image details. Finally, the perceptual loss $L_p$ measures the dissimilarity between the predicted and ground-truth enhanced images in the feature space defined by the embedding function $\phi^k(\cdot)$, which is the output of the $k$th layer in VGG-16 network [33] pre-trained on the ImageNet dataset [32]:

$$L_p = \sum_{k=2,4,6} \|\phi^k(Y) - \phi^k(\hat{Y})\|_1.$$  \hspace{1cm} (13)

In this work, the informative multi-channel enhanced image $\hat{Y}_L$ is essential to achieve high-quality results. To ensure this, we apply the exposure correction loss $L_e$ not only to the final estimation $Y$, but also to $\hat{Y}_L$. For this purpose, we convert $\hat{Y}_L$ into the image domain, i.e., $\hat{Y}_L \in \mathbb{R}^{H \times W \times 3}$, using one auxiliary $1 \times 1$ convolution. Then, we compute the exposure correction loss between $Y_L$ and $\hat{Y}_L$.

#### 4. Experiments

##### 4.1. Setting

**Datasets.** For evaluation of the proposed multi-exposure image error correction network, we use two representative multi-exposure datasets: ME [2] and SICE [6]. The ME dataset [2] consists of multi-exposure scenes, where each scene contains 5 exposure-level images and their corresponding ground-truth image. ME is decomposed into 3,535 and 1,181 scenes for training and test sets, respectively. We use the training scenes for training the proposed network and the test scenes for evaluation. The SICE dataset [6] contains 589 scenes with various exposures and corresponding ground truth. For the evaluation of SICE [6], we sample 5 exposure subsets based on the middle exposure subset for each scene. Also, we divide 589 scenes into 472 training and 117 test scenes. We use Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity (SSIM) for the quantitative evaluation in all experiments.

**Implementation details.** We resize input images to $384 \times 384$ and train the proposed network with a batch size of 2 using a single NVIDIA RTX A6000 GPU. We set 20 epochs for the ME dataset and 55 epochs for the SICE dataset. During training, we optimize the networks using the RMSProp optimizer and WarmUpCosine learning rate schedule with an initial learning rate of 3e-4.
<table>
<thead>
<tr>
<th>Methods</th>
<th>Under</th>
<th>Over</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PSNR</td>
<td>SSIM</td>
<td>PSNR</td>
</tr>
<tr>
<td>RetinexNet [37]</td>
<td>12.13</td>
<td>0.620</td>
<td>10.47</td>
</tr>
<tr>
<td>Zero-DCE [10]</td>
<td>14.55</td>
<td>0.588</td>
<td>10.40</td>
</tr>
<tr>
<td>GEN-LEN [20]</td>
<td>21.52</td>
<td>0.774</td>
<td>21.28</td>
</tr>
<tr>
<td>RCTN [19]</td>
<td>21.73</td>
<td>0.774</td>
<td>21.10</td>
</tr>
<tr>
<td>MSEC [2]</td>
<td>20.52</td>
<td>0.812</td>
<td>19.79</td>
</tr>
<tr>
<td>CMEC [26]</td>
<td>21.12</td>
<td>0.839</td>
<td>21.88</td>
</tr>
<tr>
<td>DRBN-ENC [14]</td>
<td>22.72</td>
<td>0.854</td>
<td>22.11</td>
</tr>
<tr>
<td>SID-ENC [14]</td>
<td>22.59</td>
<td>0.842</td>
<td>22.36</td>
</tr>
<tr>
<td>FECNet [15]</td>
<td>22.96</td>
<td>0.859</td>
<td>23.22</td>
</tr>
<tr>
<td>Ours</td>
<td>23.49</td>
<td>0.862</td>
<td>23.68</td>
</tr>
</tbody>
</table>

Table 3. Quantitative comparison on the ME [2] dataset. The best results are boldfaced, and the second ones are underlined.

![Images of results](image1)

(a) Overexposure input (b) DRBN-ENC (c) FECNet (d) Ours (e) Ground-truth

![Images of results](image2)

(f) Underexposure input (g) DRBN-ENC (h) FECNet (i) Ours (j) Ground-truth

Figure 3. Qualitative comparison on the ME [2] dataset.

<table>
<thead>
<tr>
<th>Methods</th>
<th>Under</th>
<th>Over</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PSNR</td>
<td>SSIM</td>
<td>PSNR</td>
</tr>
<tr>
<td>GEN-LEN [20]</td>
<td>21.52</td>
<td>0.774</td>
<td>21.28</td>
</tr>
<tr>
<td>RCTN [19]</td>
<td>21.73</td>
<td>0.774</td>
<td>21.10</td>
</tr>
<tr>
<td>DRBN-ENC [14]</td>
<td>21.39</td>
<td>0.768</td>
<td>20.49</td>
</tr>
<tr>
<td>FECNet [15]</td>
<td>21.46</td>
<td>0.788</td>
<td>20.33</td>
</tr>
<tr>
<td>Ours</td>
<td>22.34</td>
<td>0.773</td>
<td>21.76</td>
</tr>
</tbody>
</table>

Table 4. Quantitative comparison on the SICE [6] dataset. The best results are boldfaced, and the second ones are underlined.

4.2. Comparison with State-of-the-arts

**ME dataset.** Table 3 compares the proposed method with the existing exposure correction methods [37, 10, 20, 19, 2, 26, 14, 15] on the ME dataset. Following MSEC, Table 3 shows the PSNR and SSIM scores for underexposure and overexposure scenes and their average scores. For comparison, we obtain the results of the existing methods [37, 10, 2, 26, 14, 15] from [2, 15, 26]. Also, we obtain the results of methods [20, 19] using their published source codes and default settings. RetinexNet [37] and Zero-DCE [10] provide lower performance since they cannot consider multi-exposure examples. In contrast, the proposed network provides remarkable performance on the ME dataset. As compared with color transform-based methods [20, 19], the proposed method significantly improves the performance by exploiting luminance features, extracted from the luminance comparison module. Also, the proposed method outperforms recent state-of-the-art methods ENC [14] and FECNet [15], which are designed for multi-exposure scenes, for example, by a margin of 0.45 in terms
of average PSNR against FECNet [15]. Figure 3 qualitatively compares the proposed method with ENC and FECNet on the ME dataset. We see that the proposed method corrects both overexposure and underexposure scenes more faithfully than ENC and FECNet.

**SICE dataset.** Table 4 shows PSNR and SSIM scores for the proposed method and the existing methods [20, 19, 14, 15] on the SICE dataset. For comparison, we obtain the results of [20, 19, 14, 15] using their published source codes and default settings. The proposed method surpasses both color transform-based methods [20, 19] and multi-exposure correction methods [14, 15]. Figure 4 illustrates exposure correction results of ENC, FECNet, and the proposed method. In these examples, the proposed method enhances both overexposure and underexposure images to have more similar color tones to their ground truth than ENC and FECNet.

### 4.3. Ablation Study

We present a comprehensive analysis of the efficacy of our proposed components: luminance comparison module (LCM); post-processing module (PPM); the number of transformation functions ($n_t$); loss functions. All ablation studies are performed on the ME [2] dataset.

#### Luminance comparison module. We conduct experiments in three ways to analyze the contribution of LCM: First,
Methods | Under | Over | Average
--- | --- | --- | ---
L_c | PSNR 23.11 SSIM 0.848 | PSNR 23.03 SSIM 0.855 | PSNR 23.07 SSIM 0.852
L_c + L_p | PSNR 23.30 SSIM 0.849 | PSNR 23.35 SSIM 0.865 | PSNR 23.33 SSIM 0.859
L_c + L_p + L_f | **23.49** SSIM **0.862** | **23.68** SSIM **0.872** | **23.57** SSIM **0.869**

Table 7. Results on the ME [2] dataset with different loss functions. The best results are boldfaced.

Methods | Under | Over | Average
--- | --- | --- | ---
w/o aux | PSNR 23.41 SSIM 0.858 | PSNR 23.60 SSIM 0.869 | PSNR 23.49 SSIM 0.865
All (Ours) | **23.49** SSIM **0.862** | **23.68** SSIM **0.872** | **23.57** SSIM **0.869**

Table 8. Results on the ME [2] dataset depending on additional exposure correction loss. The best results are boldfaced.

Table 9. Results for ME [2] dataset with different numbers of intensity transformation functions. The best results are boldfaced.

Table 9 presents an evaluation of the effectiveness of our color transformation method based on multi-channel representation. A comparison is made between our method and an alternative approach that uses a single intensity transformation function for each color channel. The results demonstrate that our method achieves better scores. Moreover, our method exhibits an advantage in correcting overexposure by leveraging the local structures of the input image and estimating appropriate transformation functions for these structures.

5. Conclusion

We proposed the multiple exposure correction network to perform LACT. The key insights of LACT are the critical role of luminance in correcting images with multiple exposures and the need to overcome the limitation of a single intensity transformation function. We developed the luminance comparison module to extract luminance features, which encodes the exposure information. Also, we proposed LACT, which estimates the various transformation functions from the local structure based on the luminance features. LACT converts the input image into multi-channel correction representation and the post-processing module yields the exposure correction results. Experiments on ME and SCIE validated that the proposed method significantly outperforms the existing multiple exposure correction methods.

Acknowledgements. This work was supported partly by the National Research Foundation of Korea (NRF) grants (NRF-2021R1A4A1031864, NRF-2022R1I1A3069113) and partly by Institute of Information communications Technology Planning Evaluation (IITP) grant funded by the Korea government (MSIT) (No.RS-2022-00155857, Artificial Intelligence Convergence Innovation Human Resources Development (Chungnam National University)).
References


