OSMamba: Omnidirectional Spectral Mamba with Dual-Domain Prior Generator for Exposure Correction

Supplementary Material

In this supplementary material, we present additional technical details and experimental results. Specifically:

- We further provide detailed architectural designs of amplitude mamba and phase mamba along with the formulation of the core S6 algorithm and the detailed inference process of OSMamba with Dual-Domain Prior Generator.
- We provide additional ablation studies on the scan direction as well as the amplitude and phase mamba.
- We showcase additional visual comparisons on LCDP, MSEC, and SICE datasets with highlighted regions that compare the correction results among different methods.

A. More Details of Proposed Method

A.1. More Details of Amplitude and Phase Mamba

In this section, we further elaborate on the details of amplitude mamba and phase mamba. The key insight of our method lies in utilizing Omnidirectional Spectral Scanning (OS-Scan) to convert 2D amplitude and phase spectrum into 1D sequences while preserving correlations in all directions. These sequences are then modeled and modulated through the S6 mechanism. Omnidirectional Spectral Merging (OS-Merge) reconstructs the modulated sequences back into a 2D spectrum. The specific process is as follows: Initially, the amplitude spectrum or phase spectrum is processed through a linear projection followed by depth-wise convolution and SiLU activation function to enhance local feature representation. The channels of the spectrum are then divided into four equal parts. Each part is then processed using one of four distinct scanning orders in the OS-Scan. Specifically, these frequency segments are transformed into sequences along predefined traversal paths, which are then processed in parallel through the S6 mechanism. The processed sequences are subsequently reconstructed and combined via OS-Merge to obtain the output spectrum. This design ensures that each frequency point can effectively integrate information from all other points across the spectrum, establishing comprehensive global receptive fields in the frequency domain. The formulation for both amplitude and phase mamba structures is defined as:

$$\begin{aligned} \mathbf{X}' &= \text{SiLU}(\text{DWConv}(\text{Linear}(\mathbf{X}_{\text{in}}))), \\ \mathbf{X}_{\text{out}} &= \text{LN}(\text{OS-Merge}(\text{S6}(\text{OS-Scan}(\mathbf{X}')))), \end{aligned} \tag{1}$$

where $\mathbf{X}_{\text{in}}, \mathbf{X}_{\text{out}} \in \mathbb{R}^{H\times(\frac{W}{2}+1)\times D}$, with H, W, and D representing the height, width, and number of channels in the frequency domain, respectively.

Algorithm 1 S6 mechanism [1]

Input: $x : (\text{HW}, \text{C})$
Output: $y : (\text{HW}, \text{C})$
$\mathbf{A} : (\text{C}, \text{N}) \leftarrow \mathbf{Parameter}_A$
$\mathbf{B} : (\text{HW}, \text{N}) \leftarrow \mathbf{Linear}_B(x)$
$\mathbf{C} : (\text{HW}, \text{N}) \leftarrow \mathbf{Linear}_C(x)$
$\boldsymbol{\Delta} : (\text{HW}, \text{C}) \leftarrow \log(1 + \exp(\mathbf{Linear}_{\Delta}(x) + \mathbf{Parameter}_{\Delta}))$
$\overline{\mathbf{A}} : (\text{HW}, \text{C}, \text{N}) \leftarrow \exp(\boldsymbol{\Delta} \otimes \mathbf{A})$
$\overline{\mathbf{B}} : (\text{HW}, \text{C}, \text{N}) \leftarrow \boldsymbol{\Delta} \otimes \mathbf{B}$
$y \leftarrow \text{SSM}(\overline{\mathbf{A}}, \overline{\mathbf{B}}, \mathbf{C})(x)$
return y

A.2. More Details of S6 mechanism

In this section, we further elaborate on the technical details of the S6 mechanism [1], which serves as the core component in our amplitude and phase mamba designs. As shown in Algorithm 1, the S6 mechanism introduces a state dimension N and processes the input tensor $x \in \mathbb{R}^{HW\times C}$ through a state space approach. The parameters $\mathbf{B}$, $\mathbf{C}$, and $\boldsymbol{\Delta}$ are computed dynamically based on the input sequence, enabling the mechanism to adaptively model diverse temporal patterns. The state space formulation of S6 brings two key advantages: the computational complexity scales linearly with the sequence length, making it highly efficient for processing long sequences; moreover, its recursive state updates enable effective modeling of dependencies between any two positions in the sequence, regardless of their distance. These properties make S6 a powerful tool for 1D sequence modeling. Through the OS-Scan operation, both the amplitude spectrum and phase spectrum are converted into 1D sequences, allowing us to leverage these advantageous properties of the S6 mechanism. The efficient sequence modeling capability of S6 proves especially beneficial for capturing and preserving long-range spectral relationships in both amplitude and phase domains.

Algorithm 2 Inference Process of OSMamba

Input: Exposure-error images $\mathbf{I}_{\text{error}}$.
Output: Corrected images $\mathbf{I}_{\text{corrected}}$.
$\hat{\mathbf{Z}}_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$
$\mathbf{D} = \text{DDPE}^*(\mathbf{I}_{\text{error}})$
for $t = T$ to 1 **do**
$\hat{\mathbf{Z}}_{t-1} = \frac{1}{\sqrt{\alpha_t}}\left(\hat{\mathbf{Z}}_t - \frac{1-\alpha_t}{\sqrt{1-\bar{\alpha}_t}}\boldsymbol{\epsilon}_{\boldsymbol{\Theta}}(\hat{\mathbf{Z}}_t, t, \mathbf{D})\right)$
end for
Return $\mathbf{I}_{\text{corrected}} = \text{UNet}(\mathbf{I}_{\text{error}}, \hat{\mathbf{Z}}_0)$.

Setting	Model Type	PSNR	SSIM
(a)	Convolution	24.32	0.8760
(b)	Attention	24.37	0.8765
OSMamba	**OS-SSM**	**24.53**	**0.8773**

Table 1. **Ablation Study on the Amplitude and Phase Mamba.** The Amplitude and Phase Mamba is replaced with either a Convolution block (two ReLU-activated 1×1 convolutions) [2] or an Attention [3] for comparison.

A.3. More Details of Inference Process

During inference, we leverage the Dual-Domain Prior Generator (DDPG), which consists of a Dual-Domain Prior Extractor* (DDPE*) and a denoising network $\boldsymbol{\epsilon}_{\boldsymbol{\Theta}}$, to generate diffusion prior. The process begins by sampling a random Gaussian noise $\hat{\mathbf{Z}}_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ as the starting point. The DDPE* processes the image with exposure-error $\mathbf{I}_{\text{error}}$ in both spatial and frequency domains to extract comprehensive prior information $\mathbf{D}$, which serves as the condition for our diffusion model. This sampling strategy effectively eliminates the dependency on $\mathbf{I}_{\text{gt}}$. Through T iterative denoising steps, the model $\boldsymbol{\epsilon}_{\boldsymbol{\Theta}}$ progressively refines the initial noise $\hat{\mathbf{Z}}_T$ into $\hat{\mathbf{Z}}_0$ following the degeneration-free distribution of priors extracted from ground truth image, where each step is conditioned on $\mathbf{D}$. Each step follows a learned reverse diffusion process, where the noise is gradually removed according to the schedule defined by α_t. This generative prior is then injected into the Omnidirectional Spectral SSB (OS-SSB) blocks of UNet, where it further guides the network to recover fine details in severely degraded regions that remain challenging to restore after OS-SSM modulation, ultimately producing the corrected image $\mathbf{I}_{\text{corrected}}$.

B. More Ablation Experiment

To validate the effectiveness of the Amplitude and Phase Mamba, we conduct experiments by replacing them with alternative components, as shown in Table 1. In setting (a), when replacing the Amplitude and Phase Mamba with Convolution blocks [2], the PSNR drops by 0.21dB compared to OSMamba. This degradation can be attributed to the limited receptive field of convolutions, which fails to establish effective dependencies between spectral components. In setting (b), substituting the Mamba modules with Attention [3] leads to a 0.16dB decrease in PSNR, demonstrating that our proposed OS-SSM is more effective than the classical linear attention mechanism for modeling spectral relationships.

In setting (a) of Table 2, we extend the four-directional OS-Scan with their reverse directions, obtaining eight-directional scanning. However, the experimental results show that adding reverse scanning directions yields no improvements in PSNR. This indicates that the forward scanning directions are sufficient to capture the essential spectral dependencies, while the reverse directions provide redundant information. Therefore, we adopt the more efficient four-directional OS-Scan without reverse scanning in our final model.

Setting	Scan Type	PSNR	SSIM
(a)	OS-Scan w/ Reverse Direction	24.51	**0.8775**
OSMamba	**OS-Scan w/o Reverse Direction**	**24.53**	0.8773

Table 2. **Ablation study on scan direction.** Comparison between OS-Scan with and without reverse direction.

C. More Visual Comparisons

Figures 1, 2, and 3 show more visual comparison results on the mixed-exposure error dataset LCDP, multi-exposure error dataset MSEC and multi-exposure error dataset SICE, respectively. To facilitate comparison, we highlight regions with significant differences using red boxes, where our method demonstrates superior performance in preserving details and correcting exposure errors.

References

[1] Albert Gu and Tri Dao. Mamba: Linear-time sequence modeling with selective state spaces. *arXiv preprint arXiv:2312.00752*, 2023. 1

[2] Jie Huang, Yajing Liu, Feng Zhao, Keyu Yan, Jinghao Zhang, Yukun Huang, Man Zhou, and Zhiwei Xiong. Deep fourier-based exposure correction network with spatial-frequency interaction. In *European Conference on Computer Vision*, pages 163–180. Springer, 2022. 2

[3] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, and Ming-Hsuan Yang. Restormer: Efficient transformer for high-resolution image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5728–5739, 2022. 2

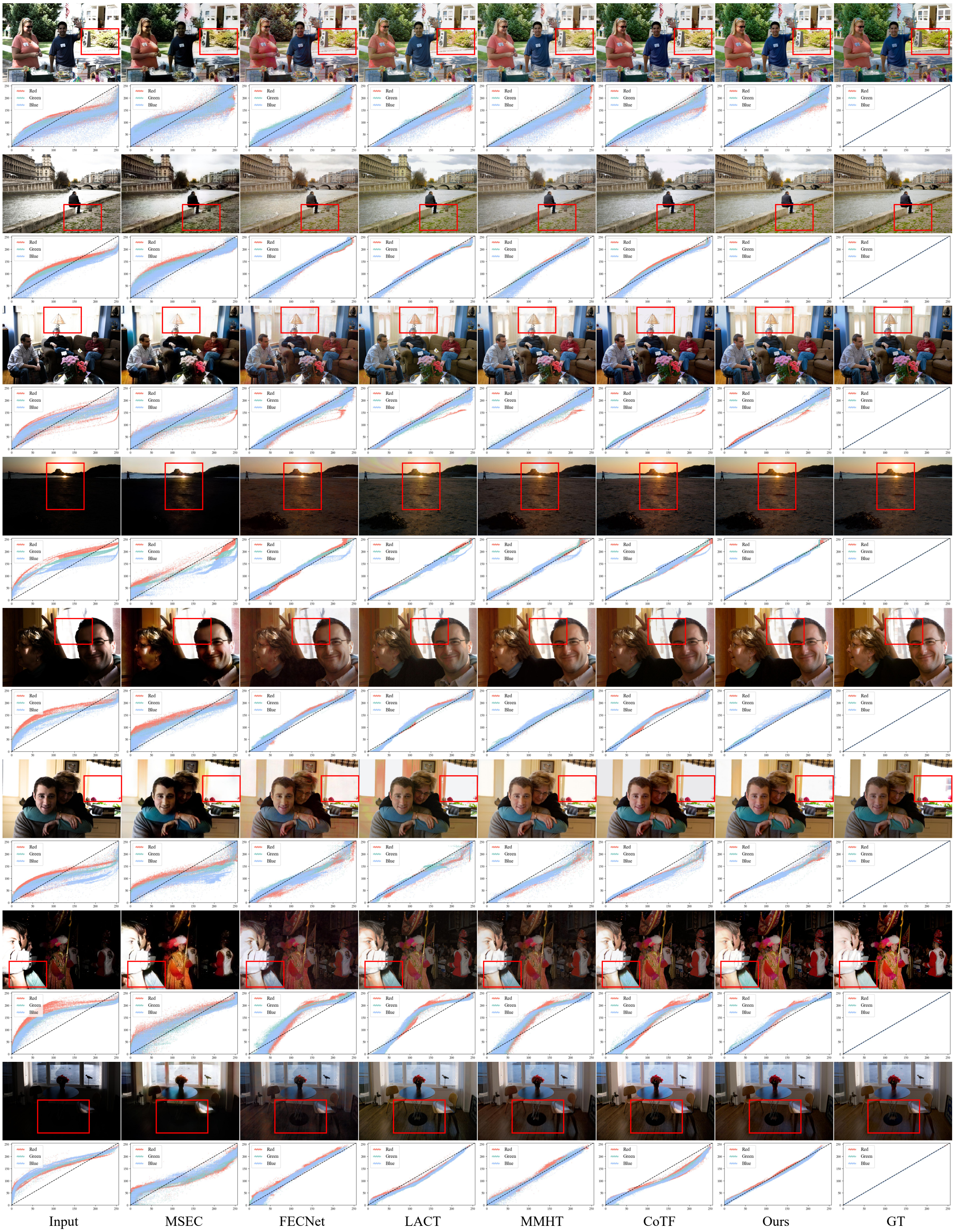

Figure 1. **More Visual comparison of our method against state-of-the-art approaches on the mixed-exposure error dataset LCDP.**

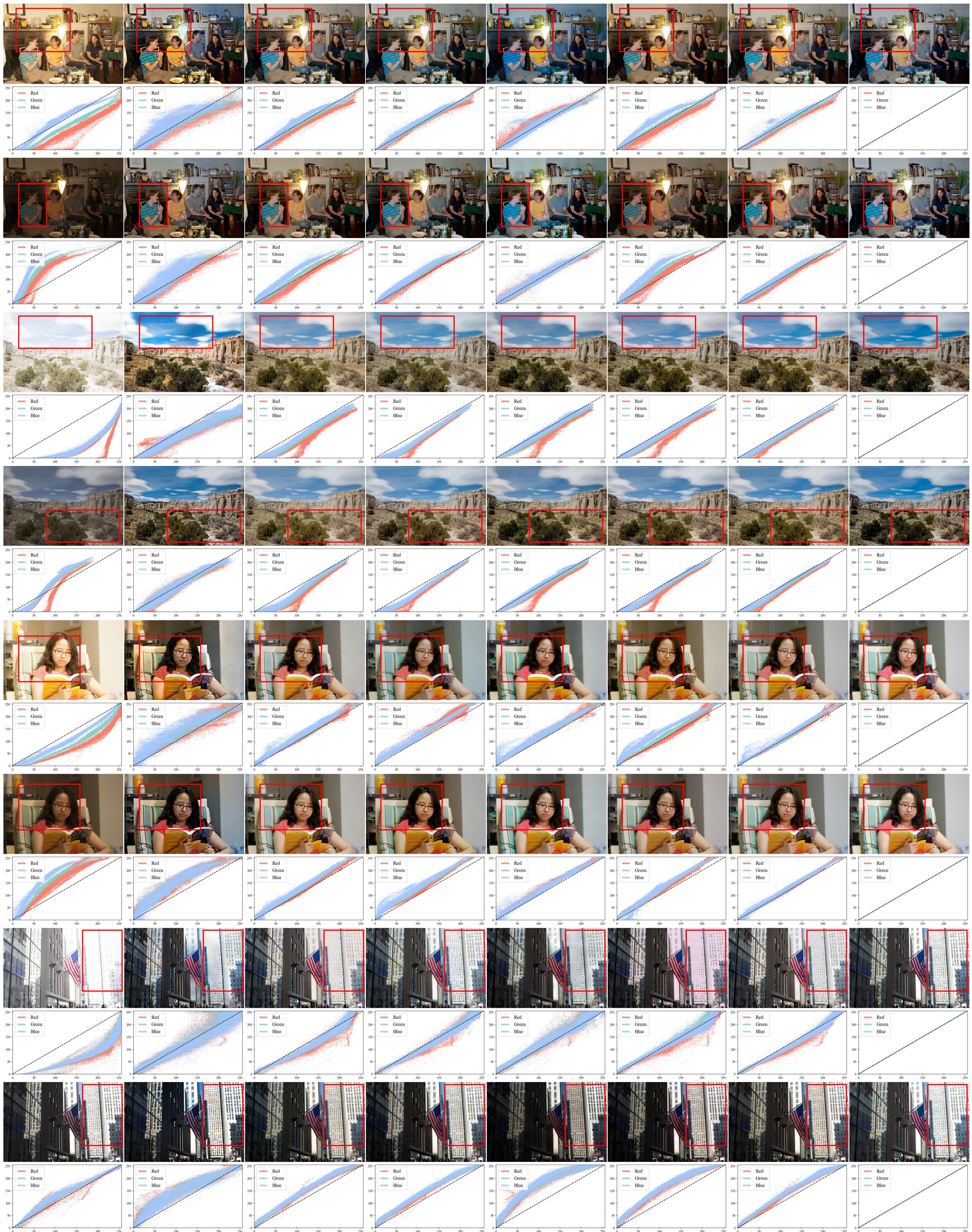

Figure 2. **More Visual comparison of our method against state-of-the-art approaches on the multi-exposure error dataset MSEC.**

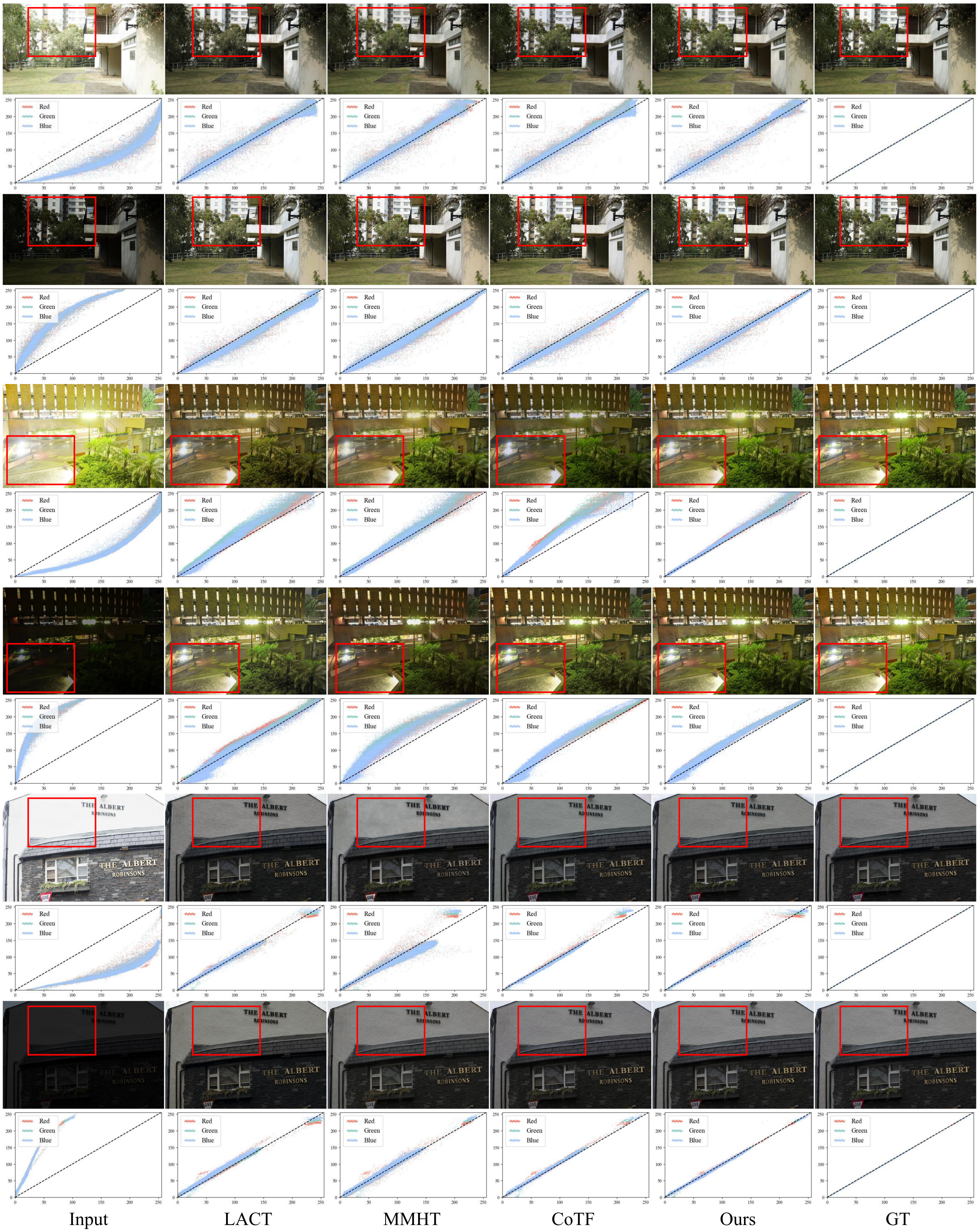

Figure 3. **More Visual comparison of our method against state-of-the-art approaches on the multi-exposure error dataset SICE.**