Dynamic Spatial Aggregation Network for Joint Denoising and HDR Imaging

Jun Xiao, Cuixin Yang, Cong Zhang, Yakun Ju, Kin-Man Lam
Department of Electrical and Electronic Engineering, The Hong Kong Polytechnic University

***Abstract*—Imaging systems often produce underexposed or overexposed images due to their limited capacity to capture the complete range of natural illumination. To tackle this limitation, high dynamic range (HDR) imaging techniques have demonstrated effectiveness in recent years. However, most existing methods focus on clean images with small-scale object motions, which are impractical for real-world scenarios. In this paper, we consider a realistic scene where the captured images exhibit spatially varying noise and experience large-scale object motions. To handle these issues, we propose a dynamic network for jointly performing denoising and HDR imaging.**

Key Words: HDR imaging, Image Processing, Image Enhancement

I. INTRODUCTION

In recent years, high dynamic range (HDR) imaging methods have significantly improved the visual quality of images, by using multiple low dynamic range (LDR) images with different exposure times. However, these methods solely focus on clean images with small object motions, rendering them ill-suited for real-world situations and leading to unsatisfactory performance. In this paper, we investigate mobile images, where substantial noise and large-scale object motions pose significant challenges to HDR imaging. To address these issues, we propose a dynamic network for jointly denoising and HDR imaging.

II. METHODOLOGY

Our method takes a set of LDR images captured with three different exposures, denoted by $\{I_L^{(1)}, I_L^{(2)}, I_L^{(3)}\}$, as input and generates a high-quality HDR image $I_H^{(2)}$. First, we linearize the input LDR images as follows:

$$\hat{I}_L^{(i)} = \frac{I_L^{(i)}}{t_i}, \quad i = 1,2,3, \tag{1}$$

where t_i is the exposure time of the i-th LDR image, and $\hat{I}_L^{(i)}$ denotes its linearized LDR image. Then, we concatenate the images to form the input of the model, i.e., $X_i = [I_L^{(i)}, \hat{I}_L^{(i)}]$.

Fig.1 demonstrates the overall structure of the proposed method, which consists of two important blocks, i.e., the conditional dynamic filtering block and the reconstruction block. The dynamic filtering block takes the non-reference LDR features and the reference features as input and adaptively generates convolutional kernels via tensor decomposition, i.e., $\mathrm{W}(F') = W_0 + P\Phi(F')Q^T$, where $F' = [F_i, F_2]$ are the concatenate features, and $i = 1, 3$. $\mathrm{W}(F')$ denotes the dynamic kernel, and W_0 represents a fixed kernel. $\Phi(F')$ is the low-dimensional representation of F'. P and Q^T denote the projected matrices. The resulting dynamic kernels can effectively mitigate adverse impacts of noise and adaptively extract features for HDR imaging. Then, we concatenate the outputs of the conditional dynamic filtering blocks with the reference features and fuse them using a convolutional layer. The fused features are forwarded to sequentially cascaded reconstruction blocks for reconstruction. We apply the skip connection to integrate the reference features with the output of the reconstruction blocks and generate the final HDR images $I_H^{(2)}$.

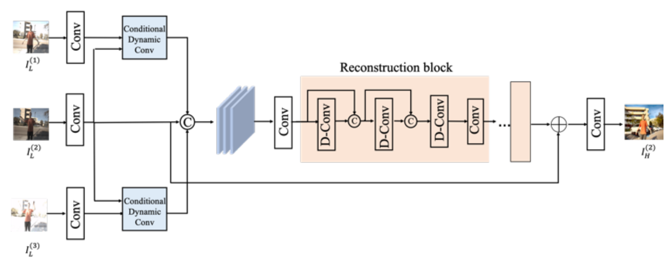

Figure 1 The overall structure of the proposed dynamic HDR network.

III. PRIMARY RESULTS

We compare our proposed method with three state-of-the-art deep HDR methods [1-3]. The peak-signal-to-noise ratio (PSNR) and structural similarity index (SSIM) are adopted to evaluate their performance on a public HDR dataset [4] in the linear domain (i.e., PSNR-ℓ and SSIM-ℓ) and tone-mapped domain (i.e., PSNR-μ and SSIM-μ). The results are tabulated in Table 1. It is observed that our proposed method surpasses the second-best method (i.e., AHDRNet) by up to 0.65dB and 1.31 dB, in terms of PSNR-μ and PSNR-ℓ, respectively, in the raw domain. This reveals that our proposed method is more effective than other methods in the joint task of denoising and HDR imaging.

TABLE I. QUANTITATIVE COMPARISON OF THE PROPOSED METHOD WITH OTHER STATE-OF-THE-ART HDR IMAGING METHODS. THE BEST RESULTS ARE HIGHLIGHTED IN BOLD.

Methods	PSNR-μ	SSIM-μ	PSNR-ℓ	SSIM-ℓ
DeepHDR [1]	37.94	0.9559	35.29	0.9869
AHDRNet [2]	38.16	0.9571	35.49	0.9872
NHDRRNet [3]	37.57	0.9517	34.54	0.9850
Ours	**38.81**	**0.9600**	**36.80**	**0.9906**

REFERENCES

[1] Shangzhe Wu, et al. "Deep high dynamic range imaging with large foreground motions." ECCV, 2018.
[2] Qingsen Yan, et al. "Attention-guided network for ghost-free high dynamic range imaging." CVPR, 2019.
[3] Qingsen Yan, et al. "Deep HDR imaging via a non-local network." IEEE Transactions on Image Processing 29 (2020): 4308-4322
[4] Shuaizheng Liu, et al. "Joint HDR Denoising and Fusion: A Real-World Mobile HDR Image Dataset." CVPR, 2023.