Reconstructing HDR Image from a Single Filtered LDR Image Base on a Deep HDR Merger Network

Bin Liang*
MRAD of Beijing Institute of Technology

Dongdong Weng†
MRAD of Beijing Institute of Technology
AICFVE of Beijing Film Academy

Yihua Bao‡
AICFVE of Beijing Film Academy

Ziqi Tu§
MRAD of Beijing Institute of Technology

Le Luo¶
MRAD of Beijing Institute of Technology

ABSTRACT

In this paper, a novel Deep HDR Merger network, which is called MergeNet, is proposed to reconstruct a HDR image from a single filtered LDR image. Filtered images are adopted as input since they contain more dynamic range than traditional ones. By learning the correlation between filtered LDR images and HDR images, the MergeNet successfully achieves HDR reconstruction of filtered images. We used five evaluation methods to make qualitative and quantitative comparisons to show that our method produced excellent results. Experimental results show that the proposed method performs favorably against state-of-the-art HDR image reconstruction methods.

Index Terms: Computing methodologies—Computer graphics—Image processing;

1 INTRODUCTION

High dynamic range (HDR) images have been applied in many fields such as realistic photography editing, physically-based rendering, games, and film. Debevec [1] proposed using high dynamic range imaging (he called 'light probe') for rendering virtual scenes. Classic algorithms [1,2] are to merge multi-exposure LDR images to recover an HDR image. But what if there's only a single exposure? These algorithms [5,7,9,11] can be implemented by traditional manual modeling. In recent years, the convolution neural networks (CNNs) also can be utilized for HDR reconstruction. Eilertsen et al. [3] used U-Net architecture that operates to **directly** create a logarithmic HDR image from an LDR image. Endo et al. [4] used an improved U-Net architecture that **indirectly** predicts multi-exposure LDR images from a single image and then uses standard merging algorithms to generate an HDR content. These two methods essentially treat a single color image as three monochrome images in the same exposure state. In this paper, a specific optical filter, which changes the transmittance of three channels (see Fig. 1 (top right)), is placed in the front of camera lens to enhance the dynamic range of the image [8, 12]. In this paper, a new modify CNNs called HDR Merger Network (MergeNet, see Fig. 1) is proposed. we input a single filtered image and obtain one logarithmic HDR image (direct mode) and seven multi-exposure LDR images (indirect mode). The two types of images are processed separately and merged into a final HDR image.

2 HYBRID HDR MERGER NETWORK

Step 1: Creating LDR images We use HDR data as the scene radiance and construct the LDR data using the virtual camera imaging formula (1):

$$L_{i,j} = f(H_i \Delta t_j), \Delta t_j = \frac{1}{\tau^{T/2}}, ..., \frac{1}{\tau^2}, \frac{1}{\tau}, 1, \tau, \tau^2, ..., \tau^{T/2} \quad (1)$$

Where H_i denotes the $i.th$ HDR image irradiance values, Δt_j denotes the $j.th$ exposure duration, f denotes the camera response function (CRF) from Grossberg and Nayar's Database of Response Functions (DoRF) [6], $L_{i,j}$ denotes LDR image pixel values for each pixel i and exposure time index j. Here, We use $\tau = \sqrt{2}, T = 8$.

Step 2: Constructing MergeNet Network The MergeNet is built on a U-net [3, 4] network, see Fig. 1. In particular, the input is a filtered image and the output is a logarithmic HDR image and $N-1$ multi-exposure LDR images.

Step 3: Loss Function Take up-exposure model as an example, the down-exposure model just needs to change the order, namely,

$$\mathscr{L}_{j,i} = \left\| \{H_j, I_{j,i+1\to j,i+N}\} \otimes Q_i - M_i \circ G(I^F_{j,i}, \theta) \right\|_1 \quad (2)$$

Where, H_j denotes a logarithmic HDR image. $I_{j,i+1\to j,i+N}$ denotes a $min\{N-1, T-i\} \times W \times H \times c$ tensor. $\otimes$ and O_i are a concatenation operator and a $\min\{i+N-T-1, 0\} \times W \times H \times c$ zero tensor. M_i and $\circ$ are a mask tensor and an element-wise product operator. Each element of M_i is one if the index of the first dimension of $\{H_j, I_{i+1\to i+N}\}$ is less than $T+2-i$ and zero otherwise. $G(I^F_{j,i}, \theta)$ denotes an output tensor when inputting the $i.th$ filtered image $I^F_{j,i}$ and θ is the network model parameters.

Step 4: Merging HDR Images Feeding a filtered LDR image to the MergeNet, the up-/down-exposure network output two logarithmic HDR images and $2(N-1)$ multi-exposure LDR images which sort from the darkest to the brightest. The final HDR image is merged as follow, here we set $\alpha = 0.6$:

$$H^{log}_j = 10^{(H^{up}_j + H^{down}_j)/2}, H^{merge}_j = Debevec(LDRs, times) \quad (3)$$

$$H^{final}_j = (1-\alpha) H^{log}_j + \alpha \frac{H^{merge}_j}{max(H^{merge}_j)} \times max(H^{log}_j) \quad (4)$$

3 EXPERIMENTS

Dataset: To train our network, we constructed a training dataset by collecting HDR datasets from DML-HDR online[1] and funt-HDR[2]. In this work, we use 43,200 triplets to train and use 4800 triplets to test. Note that test data and input of resulting images both do not appear in the training data.

Evaluation Methods and Metrics: In this paper, we compared seven iTMO methods including Landis (LEO) [9], Akyz et al. [5], Kovaleski and Oliveira [11] (KOEO), Huo et al. (HEO) [7], HDR-CNN [3], DRTMO [4] and our MergeNet. We adopt five evaluation metrics including Mean Square Error (MSE), Peak Signal Noise Ratio (PSNR), Structural Similarity (SSIM), Feature Similarity (FSIM) and HDR-VDP-2 [10] to compare images performances. Note that

*e-mail: 18203416228@163.com
†e-mail: crgj@bit.edu.cn
‡e-mail: baoyihua@bfa.edu.cn
§e-mail: tzq1991@bit.edu.cn
¶e-mail: l_luo@bit.edu.cn

[1] http://dml.ece.ubc.ca/data/DML-HDR/
[2] https://www2.cs.sfu.ca/ colour/data/

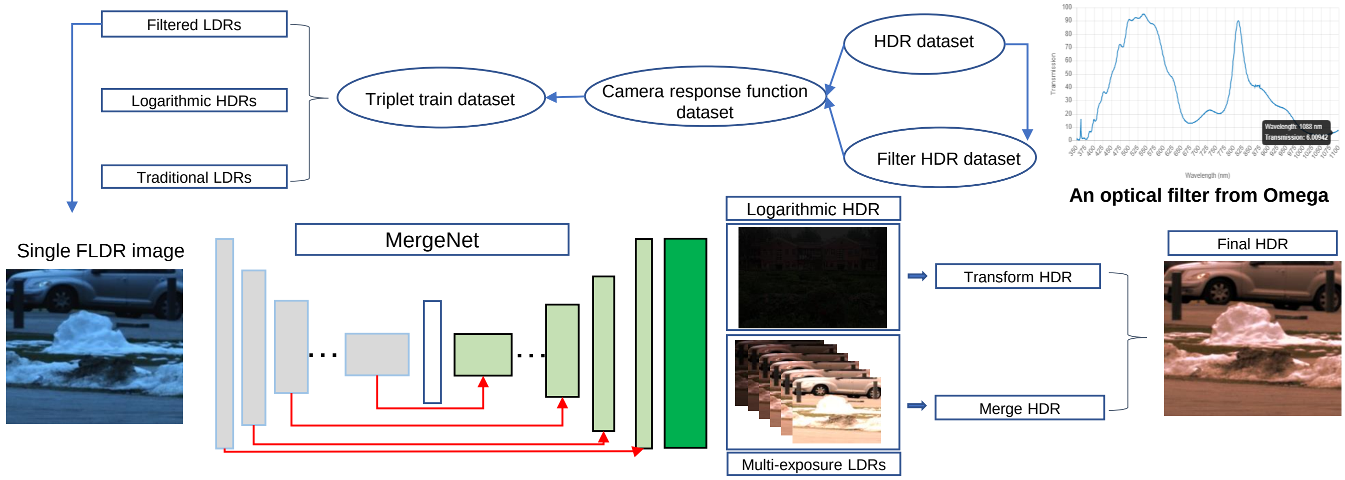

Figure 1: **An overview of the proposed method** consists of generating training dataset and reconstructing HDR phases. In generating phase, the triplet pairs are produced from HDR datasets by simulating cameras. Next, we forecast and merge a final HDR image by MergeNet network.

Table 1: Average values of the three exposure states (under-/normal-/over-exposure) for all methods. Bold values indicate the best value.

Method	MSE	PSNR	SSIM	FSIM	HDR-VDP-2
		Under-exposure			
LEO	2032.4	18.67	0.6035	0.8583	56.07
AEO	623.0	23.46	0.8669	0.9399	62.73
KOEO	567.3	22.65	0.8414	0.9455	62.51
HEO	1086.8	18.91	0.7141	0.9104	57.83
HDRCNN	428.1	24.06	0.8756	0.9554	65.77
DRTMO	528.2	23.13	0.8716	0.9535	67.20
Ours	**213.3**	**27.81**	**0.9476**	**0.9688**	**67.91**
		Normal-exposure			
LEO	3464.6	14.48	0.3959	0.7268	54.73
AEO	2564.1	17.43	0.7945	0.8758	58.84
KOEO	1172.0	20.20	0.8235	0.9021	57.19
HEO	1701.8	18.46	0.7915	0.8948	58.27
HDRCNN	859.6	23.14	0.8794	0.9302	63.02
DRTMO	809.1	21.90	0.8780	0.9307	62.83
Ours	**514.1**	**26.19**	**0.9215**	**0.9558**	**64.23**
		Over-exposure			
LEO	4982.6	12.20	0.3926	0.6660	53.14
AEO	7035.5	11.72	0.6366	0.7630	54.03
KOEO	12092.5	8.360	0.0280	0.5073	53.56
HEO	5549.2	13.03	0.6899	0.7994	54.76
HDRCNN	1225.6	20.46	**0.8353**	0.8942	59.75
DRTMO	1640.6	18.18	0.7867	0.8657	54.03
Ours	**1187.7**	**21.53**	0.8344	**0.8982**	**61.14**

we convert all HDR predicted images from RGB space to HSL, which can better compare the brightness information.

Results: Table 1 show the HDR reconstruction performance of under-, over-, and normal-exposure images respectively. Compared with the others, almost all of our method's evaluation scores are better than others. In particular, only one indicator (SSIM) is slightly lower but the gap is not very obvious.

4 CONCLUSION

This paper proposes a novel deep HDR Merger network for HDR image reconstruction from a single filtered LDR image. The results show that the MergeNet method is more advanced than other existing HDR reconstruction methods. In the future work, we will choose a suitable filter to generate real data, and train the network together with the virtual data. Further, it will be possible to restore any filtered image to the original image and even to achieve a more spectacular artistic effect.

ACKNOWLEDGMENTS

This work was supported by the National Key Research and Development Program of China (No.2017YFB1002805) and the National Natural Science Foundation of China (No.U1605254).

REFERENCES

[1] P. Debevec. Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography. In *ACM SIGGRAPH 2008 classes*, p. 32. ACM, 2008.

[2] P. E. Debevec. Recovering high dynamic range radiance maps from images. In *Conference on Computer Graphics & Interactive Techniques*, 1997.

[3] G. Eilertsen, J. Kronander, G. Denes, R. K. Mantiuk, and J. Unger. Hdr image reconstruction from a single exposure using deep cnns. *ACM Transactions on Graphics (TOG)*, 36(6):178, 2017.

[4] Y. Endo, Y. Kanamori, and J. Mitani. Deep reverse tone mapping. *ACM Trans. Graph.*, 36(6):177–1, 2017.

[5] R. Fleming, B. E. Riecke, and E. Reinhard. Do hdr displays support ldr content?: a psychophysical evaluation. In *Acm Siggraph*, 2007.

[6] M. D. Grossberg and S. K. Nayar. What is the space of camera response functions? In *2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings.*, vol. 2, pp. II–602. Citeseer, 2003.

[7] Y. Huo, Y. Fan, D. Le, and V. Brost. Physiological inverse tone mapping based on retina response. *Visual Computer*, 30(5):507–517, 2014.

[8] G. Krawczyk, M. Goesele, and H.-P. Seidel. Photometric calibration of high dynamic range cameras. 2005.

[9] H. Landis. Production-ready global illumination. *Siggraph course notes*, 16(2002):11, 2002.

[10] M. Narwaria, R. Mantiuk, M. P. Da Silva, and P. Le Callet. Hdr-vdp-2.2: a calibrated method for objective quality prediction of high-dynamic range and standard images. *Journal of Electronic Imaging*, 24(1):010501, 2015.

[11] M. M. Oliveira and R. P. Kovaleski. High-quality reverse tone mapping for a wide range of exposures. In *Graphics, Patterns & Images*, 2014.

[12] R. Shrestha, A. Mansouri, and J. Y. Hardeberg. Multispectral imaging using a stereo camera: Concept, design and assessment. *EURASIP Journal on Advances in Signal Processing*, 2011(1):57, 2011.