Image Correction via Deep Reciprocating HDR Transformation

Xin Yang[2,1⋆], Ke Xu[1,2⋆], Yibing Song[3†], Qiang Zhang[2], Xiaopeng Wei[2], Rynson W.H. Lau[1]

[1]City University of Hong Kong [2]Dalian University of Technology [3]Tencent AI Lab

https://ybsong00.github.io/cvpr18_imgcorrect/index
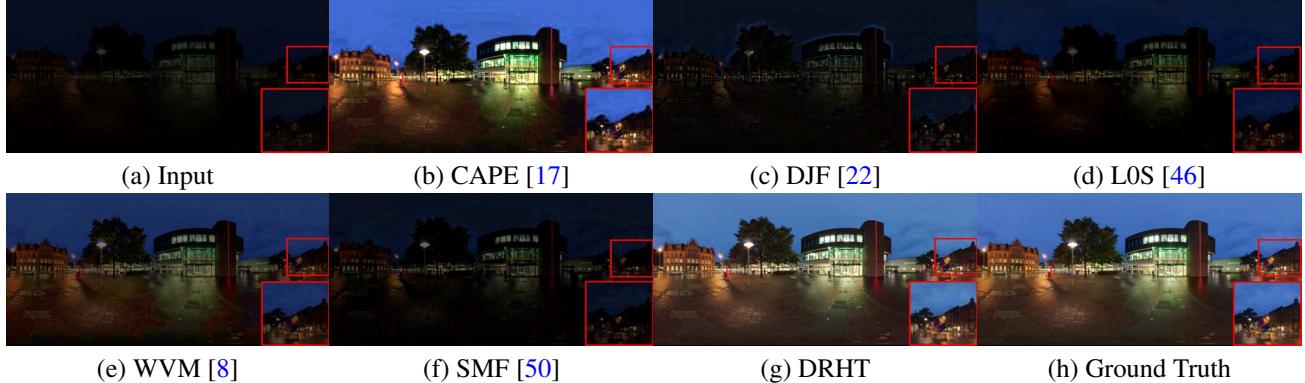
Figure 1: Image correction results on an underexposed input. Existing LDR methods have the limitation in recovering the missing details, as shown in (b)-(f). In comparison, we recover the missing LDR details in the HDR domain and preserve them through tone mapping, producing a more favorable result as shown in (g).

Abstract

Image correction aims to adjust an input image into a visually pleasing one. Existing approaches are proposed mainly from the perspective of image pixel manipulation. They are not effective to recover the details in the under/over exposed regions. In this paper, we revisit the image formation procedure and notice that the missing details in these regions exist in the corresponding high dynamic range (HDR) data. These details are well perceived by the human eyes but diminished in the low dynamic range (LDR) domain because of the tone mapping process. Therefore, we formulate the image correction task as an HDR transformation process and propose a novel approach called Deep Reciprocating HDR Transformation (DRHT). *Given an input LDR image, we first reconstruct the missing details in the HDR domain. We then perform tone mapping on the predicted HDR data to generate the output LDR image with the recovered details. To this end, we propose a united framework consisting of two CNNs for HDR reconstruction and tone mapping. They are integrated end-to-end for joint training and prediction. Experiments on the standard benchmarks demonstrate that the proposed method performs favorably against state-of-the-art image correction methods.*

1. Introduction

The image correction problem has been studied for decades. It dates back to the production of Charge-Coupled Devices (CCDs), which convert optical perception to digital signals. Due to the semiconductors used in the CCDs, there is an unknown nonlinearity existed between the scene radiance and the pixel values in the image. This nonlinearity is usually modeled by gamma correction, which has resulted in a series of image correction methods. These methods tend to focus on image pixel balance via different approaches including histogram equalization [28], edge preserving filtering [11, 1], and CNN encoder-decoder [41]. Typically, they function as a preprocessing step for many machine vision tasks, such as optical flow estimation [3, 15], image decolorization [37, 36], image deblurring [30, 29], face stylization [39, 35] and tracking [38].

Despite the demonstrated success, existing methods have the limitation in correcting images with under/over exposure. An example is shown in Figure 1, where the state-of-the-art image correction methods fail to recover the missing details in the underexposed regions. This is because the pixel values around these regions are close to 0, and the details are diminished within them. Although different image pixel operators have been proposed for image correction,

---

the results are still unsatisfactory, due to the ill-posed nature of the problem. Thus, a question is raised if it is possible to effectively recover the missing details during the image correction process.

To answer the aforementioned question, we trace back to the image formation procedure. Today's cameras still require the photographer to carefully choose the exposure duration (Δt) and rely on the camera response functions (CRFs) to convert a natural scene (S) into an LDR image (I), which can be written as [5]:

$$I = f_{CRF}(S \times \Delta t), \qquad (1)$$

However, when an inappropriate exposure duration is chosen, the existing CRFs can neither correct the raw data in the CCDs nor the output LDR images. This causes the under/over exposure in the LDR images. Based on this observation, we propose an end-to-end framework, called *Deep Reciprocating HDR Transformation (DRHT)*, for image correction. It contains two CNN networks. The first CNN network reconstructs the missing details in the HDR domain and the second CNN network transfers the details back to the LDR domain. Through the reciprocating HDR transformation process, LDR images are corrected in the intermediate HDR domain.

Overall, the contribution in this work can be summarized as follows. We interpret image correction as the Deep Reciprocating HDR Transformation (DRHT) process. An end-to-end DRHT model is therefore proposed to address the image correction problem. To demonstrate the effectiveness of the proposed network, we have conducted extensive evaluations on the proposed network with the state-of-the-art methods, using the standard benchmarks.

2. Related Work

In this section, we discuss relevant works to our problem, including image restoration and filtering, image manipulation, and image enhancement techniques.

Image Restoration and Filtering. A variety of state-of-the-art image correction methods have been proposed. Image restoration methods improve the image quality mainly by reducing the noise via different deep network designs [19, 40, 52], low-rank sparse representation learning [21] or soft-rounding regularization [26]. Noise reduction can help improve the image quality, but cannot recover the missing details. Edge-aware image filtering techniques are also broadly studied for smoothing the images while maintaining high contrasted structures [2, 22, 33], smoothing repeated textures [23, 47, 50] or removing high contrast details [24, 54, 55]. Further operations can be done to enhance the images by strengthening the details filtered out by these methods and then adding them back. Although these

filtering methods are sensitive to the local structures, overexposed regions are usually smoothed in the output images and therefore details can hardly be recovered.

Image Manipulation. Image correction has also been done via pixel manipulation for different purposes, such as color enhancement [48] and mimicking different themes/styles [42, 43]. Son *et al.* [34] propose a tone transfer model to perform region-dependent tone shifting and scaling for artistic style enhancement. Yan *et al.* [49] exploit the image contents and semantics to learn tone adjustments made by photographers via their proposed deep network. However, these works mainly focus on manipulating the LDR images to adapt to various user preferences.

Image Enhancement. Histogram equalization is the most widely used method for image enhancement by balancing the histogram of the image. Global and local contrast adjustments are also studied in [14, 31] for enhancing the contrast and brightness. Kaufman *et al.* [17] propose a framework to apply carefully designed operators to strengthen the detected regions (*e.g.*, faces and skies), in addition to the global contrast and saturation manipulation. Fu *et al.* [8] propose a weighted variational method to jointly estimate the reflectance and illumination for color correction. Guo *et al.* [10] propose to first reconstruct and refine the illumination map from the maximum values in the RGB channels and then enhance the illumination map. Recently, Shen *et al.* [32] propose a deep network to directly learn the mapping relations of low-light and ground truth images. This method can successfully recover rich details buried in low light conditions, but it tends to increase the global illumination and generate surrealistic images.
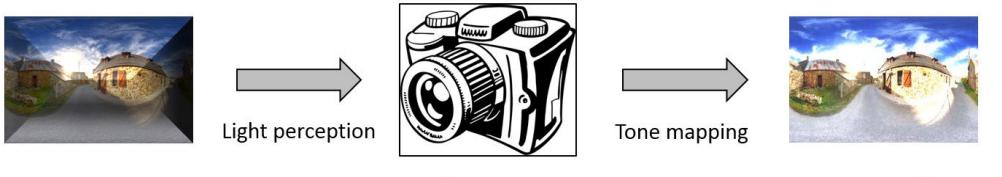
All these methods, however, cannot completely recover the missing details in the bright and dark regions. This is mainly because both their inputs and their enhancing operations are restricted to work in the LDR domain, which does not offer sufficient information to recover all the details while maintaining the global illumination.

3. Deep Reciprocating HDR Transformation
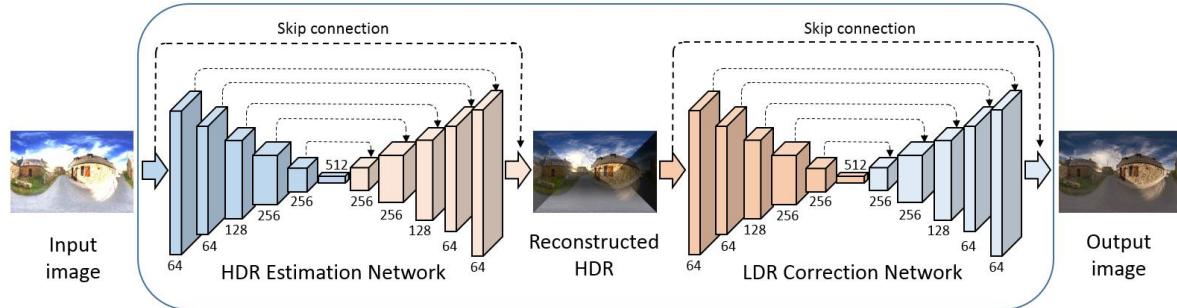
An overview of the proposed method is shown in Figure 2(b). We first illustrate our reformulation of image correction. We then show our HDR estimation network to predict HDR data given LDR input. Finally, we show that the HDR data is tone mapped into the output LDR using a LDR correction network. The details are presented as follows:

3.1. Image Correction Reformulation

Although human can well perceive the HDR data, it requires empirically configuring the camera during the imaging process. An overview of scene capturing and producing LDR is shown in Figure 2(a). However, when under

(a) Image formulation process

(b) Deep Reciprocating HDR Transformation (DRHT) pipeline

Figure 2: An overview of image formulation process and the proposed DRHT pipeline. Given an input under/over exposed LDR image, we first reconstruct the missing details in the HDR domain and map them back to the output LDR domain.

extreme lighting conditions (e.g., the camera is facing the sun), details in the natural scenes are lost during the tone mapping process. They cannot be recovered by existing image correction methods in the LDR domain.

In order to recover the degraded regions caused by under/over exposures, we trace back to the image formation procedure and formulate the correction as the Deep Reciprocating HDR Transformation process: $\hat{S} = f_1(I; \theta_1)$ and $\hat{I}^{ldr} = f_2(\hat{S}; \theta_2)$, where $\hat{S}$ and $\hat{I}^{ldr}$ represent the reconstructed HDR data and the corrected LDR image, respectively. θ_1 and θ_2 are the CNN parameters. Specifically, we propose the HDR estimation network (f_1) to first recover the details in the HDR domain and then the LDR correction network (f_2) to transfer the recovered HDR details back to the LDR domain. Images are corrected via this end-to-end DRHT process.

3.2. HDR Estimation Network

We propose a HDR estimation network to recover the missing details in the HDR domain, as explained below:

Network Architecture. Our network is based on a fully convolutional encoder-decoder network. Given an input LDR image, we encode it into a low dimensional latent representation, which is then decoded to reconstruct the HDR data. Meanwhile, we add skip connections from each encoder layer to its corresponding decoder layer. They enrich

the local details during decoding in a coarse-to-fine manner. To facilitate the training process, we also add a skip connection directly from the input LDR to the output HDR. Instead of learning to predict the whole HDR data, the HDR estimation network only needs to predict the difference between the input and output, which shares some similarity to residual learning [12]. We train this network from scratch and use batch normalization [16] and ELU [4] activation for all the convolutional layers.

Loss Function. Given an input image I, the output of this network $\hat{S} = f_1(I; \theta_1)$, and the ground truth HDR image Y, we use the Mean Square Error (MSE) as the objective function:

$$Loss_{hdr} = \frac{1}{2N} \sum_{i=1}^{N} \left\| \hat{S}_i - \alpha(Y_i)^\gamma \right\|_2^2, \qquad (2)$$

where i is the pixel index and N refers to the total number of pixels. α and γ are two constants in the nonlinear function to convert the ground truth HDR data into LDR, which is empirically found to facilitate the network convergence. We pretrain this network in advance before integrating it with the remaining modules.

3.3. LDR Correction Network

We propose a LDR correction network, which shares the same architecture as that of the HDR estimation network. It

aims to preserve the recovered details in the LDR domain, as explained below:

Loss Function. The output of the HDR estimation network $\hat{S}$ is in LDR as shown in Eq. 2. We first map it to the HDR domain via inverse gamma correction. The mapped result is denoted as $\hat{S}_{full}$. We then apply a logarithmic operation to preserve the majority of the details and feed the output to the LDR correction network. Hence, the recovered LDR image $\hat{I}^{ldr}$ through our network becomes:

$$\hat{I}^{ldr} = f_2(log(\hat{S}_{full} + \delta); \theta_2), \qquad (3)$$

where $log()$ is used to compress the full HDR domain for convergence while maintaining a relatively large range of intensity, and δ is a small constant to remove zero values. With the ground truth LDR image I^{gt}, the loss function is:

$$Loss_{ldr} = \frac{1}{2N} \sum_{i=1}^{N} (\left\| \hat{I}_i^{ldr} - I_i^{gt} \right\|_2^2 + \epsilon \left\| \hat{S}_i - \alpha(Y_i)^\gamma \right\|_2^2), \qquad (4)$$

where ϵ is a balancing parameter to control the influence of the HDR reconstruction accuracy.

Hierarchical Supervision. We train this LDR correction network together with the aforementioned HDR estimation network. We adopt this end-to-end training strategy in order to adapt our whole model to the domain reciprocating transformation. To facilitate the training process, we adopt the hierarchical supervision training strategies similar to [13]. Specifically, we start to train the encoder part and the shallowest deconv layer of the LDR correction network by freezing the learning rates of all other higher deconv layers. During training, higher deconv layers are gradually added for fine tuning while the learning rates of the encoder and shallower deconv layers will be decreased. In this way, this network can learn to transfer the HDR details to LDR domain in a coarse-to-fine manner.

3.4. Implementation Details

The proposed DRHT model is implemented under the Tensorflow framework [9] on a PC with an i7 4GHz CPU and an NVIDIA GTX 1080 GPU. The network parameters are initialized using the truncated normal initializer. We use 9×9 and 5×5 kernel sizes to generate 64-dimensional feature maps for the first two conv layers and their counterpart deconv layers for both networks, and the remaining kernel size is set to 3×3. For loss minimization, we adopt the ADAM optimizer [20] with an initial learning rate of 1e-2 for 300 epochs, and then use learning rate of 5e-5 with momentum $\beta_1 = 0.9$ and $\beta_2 = 0.998$ for another 100 epochs. α and γ in Eq. 2, and δ in Eq. 3 are set to 0.03, 0.45 and $1/255$, respectively. We also clip the gradients to avoid the gradient explosion problem. The general training takes about ten days and the test time is about 0.05s for a 256×512 image.

|			
(a) Input	DRHT (64.75)	(b) Input	DRHT (65.61)
(c) Input	DRHT (61.80)	(d) Input	DRHT (69.28)
(e) Input	DRHT (62.69)	(f) Input	DRHT (69.04)
(g) Input	DRHT (69.57)	(h) Input	DRHT (62.17)
(i) Input	DRHT (61.80)	(j) Input	DRHT (65.18)

low difference   high difference
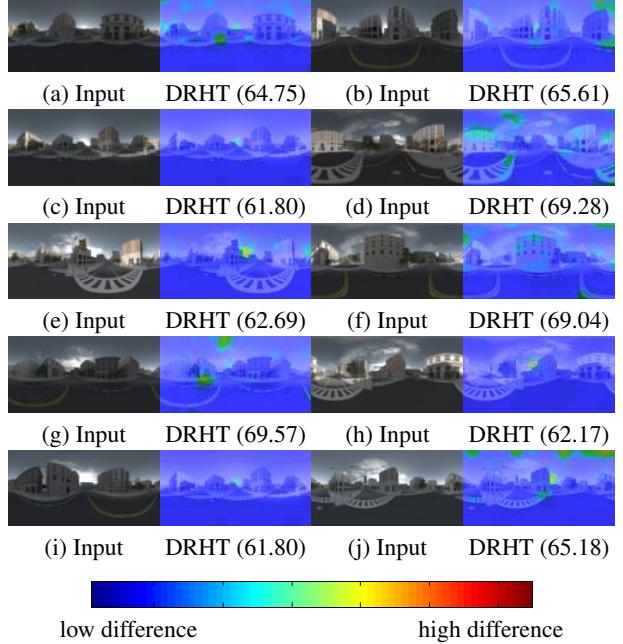
Figure 3: Internal Analysis. We compare the reconstructed HDR images with the ground truth HDR images using the HDR-VDP-2 metric. The average Q score and SSIM index on this test set are 61.51 and 0.9324, respectively.

4. Experiments

In this section, we first present the experiment setups and internal analysis on the effectiveness of the HDR estimation network. We then compare our DRHT model with the state-of-the-art image correction methods on two datasets.

4.1. Experiments Setups

Datasets. We conduct experiments on the city scene panorama dataset [51] and the Sun360 outdoor panorama dataset [45]. Specifically, since the low-resolution (64×128 pixels) city scene panorama dataset [51] contains LDR and ground truth HDR image pairs, we use the black-box Adobe Photoshop software to empirically generate ground truth LDR images with human supervision. Therefore, we use $39,198$ image pairs (i.e., the input LDR and the ground truth HDR) to train the first network and use $39,198$ triplets (i.e., the input LDR, the ground truth HDR and the ground truth LDR) to train the whole network. We use $1,672$ images from their testing set for evaluation. To adapt our models to the real images with high resolution, we use the Physically Based Rendering Technology (PBRT) [27] to generate 119 ground truth HDR scenes as well as the input and ground truth LDR images, which are then divided into $42,198$ patches for training. We also use $6,400$ images from the Sun360 outdoor panorama dataset [45] for end-to-end finetuning (i.e., ϵ in Eq. 4 is fixed as 0), as they do not
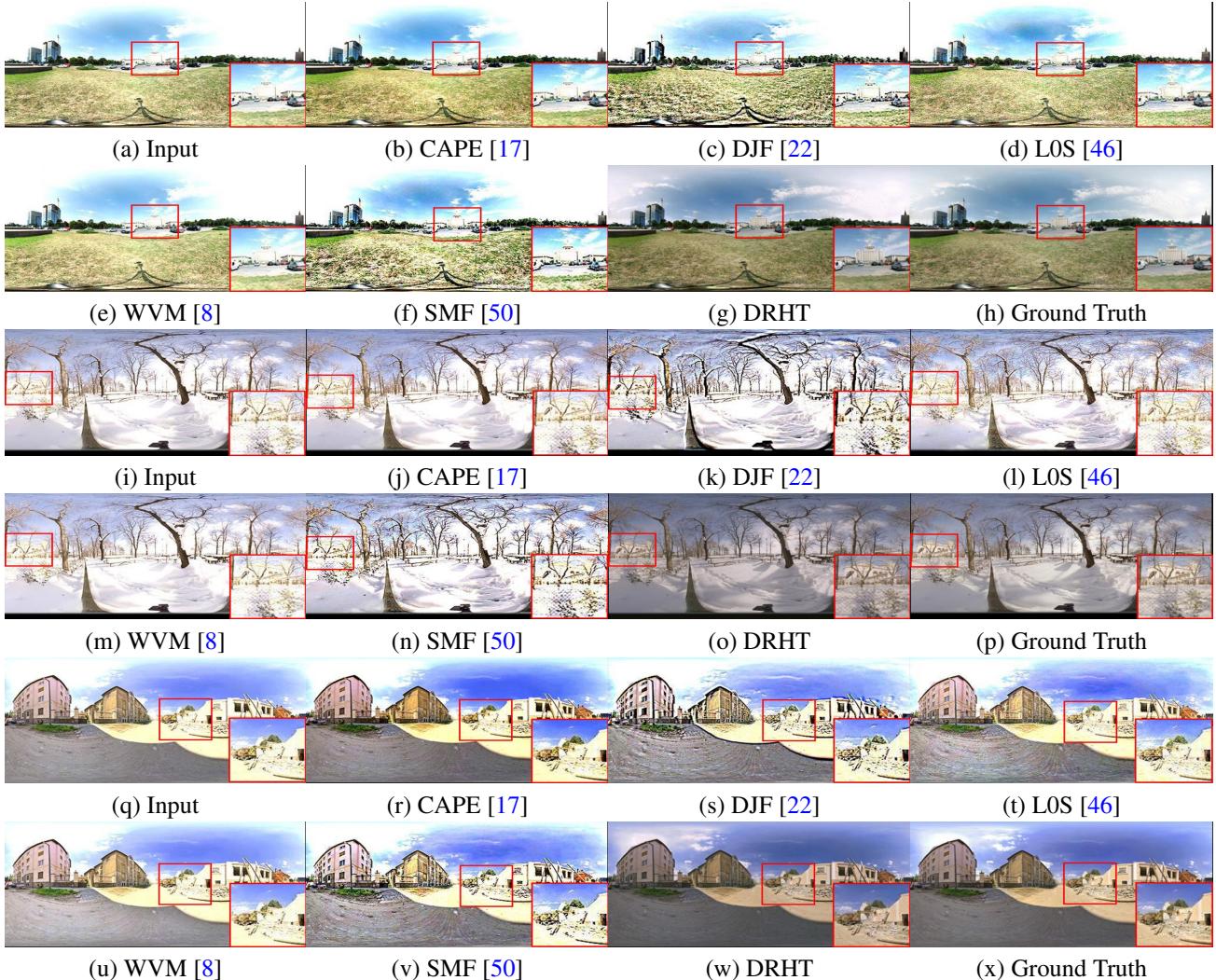
Figure 4: Visual comparison on overexposed images in the bright scenes. The proposed DRHT method can effectively recover the missing details buried in the overexposed regions compared with state-of-the-art approaches

have ground truth HDR images, and use 1, 200 images for evaluation. The input images are corrupted from the originals by adjusting the exposure (selected from the interval [-6, 3], in order not to learn the mapping between one specific exposure degree and the ground truth) and contrasts to over/under expose the visible details. We resize the images to 256×512 pixels in this dataset.

Evaluation Methods. We compare the proposed method to 5 state-of-the-art image correction methods Cape [17],WVM [8], SMF [50], L0S [46] and DJF [22] on the dataset. Among them, Cape [17] enhances the images via a comprehensive pipeline including global contrast/saturation correction, sky/face enhancement, shadow-saliency and texture enhancement. WVM [8] first decomposes the input image into reflectance and illumination

maps, and corrects the input by enhancing the illumination map. Since the enhancement operations are mostly conducted on the detail layer extracted by existing filtering methods, we further compare our results to state-of-the-art image filtering methods. Meanwhile, we compare the proposed method to two deep learning based image correction methods: Hdrcnn [6] and DrTMO [7].

Evaluation Metrics. We evaluate the performance using different metrics. When internal analyzing the HDR estimation network, we use the widely adopted HDR-VDP-2 [25] metric it reflects human perception on different images. When comparing with existing methods, we use three commonly adopted image quality metrics: PSNR, SSIM [44] and FSIM [53]. In addition, we provide the Q scores from the HDR-VDP-2 [25] metric to evaluate the image quality.

Figure 5: Visual comparison on under/over exposed images in the dark scenes. The proposed DRHT method can effectively recover the missing details in the under/over exposed regions while maintaining the global illumination.

4.2. Internal Analysis

As the proposed DRHT method first recovers the details via the HDR estimation network, we demonstrate its effectiveness in reconstructing the details in the HDR domain. We evaluate on the city scene dataset using the HDR-VDP-2 metric [25]. It generates the probability map and the Q score for each test image. The probability map indicates the difference between two images to be noticed by an observer on average. Meanwhile, the Q score predicts the quality degradation through a Mean-Opinion-score metric.

We provide some examples in Figure 3 which are from the city scene test dataset. We overlay the predicted visual difference on the generated result. The difference intensity is shown via a color bar where the low intensity is marked as blue while the high intensity is marked as red. It shows

that the proposed HDR estimation network can effectively recover the missing details on the majority of the input image. However, the limitation appears on the region where the part of sun is occluded by the building, as shown in (j). It brings high difference because the illumination contrast is high around the boundary between sun and the building. This difference is difficult to preserve in the HDR domain. The average Q score and SSIM index on this test set are 61.51 and 0.9324, respectively. They indicate that the synthesized HDR data through our HDR estimation network is close to the ground truth HDR data.

4.3. Comparison with State-of-the-arts

We compare the proposed DRHT method with state-of-the-art image correction methods on the standard bench-

Methods	City Scene dataset				Sun360 Outdoor dataset			
	PSNR	SSIM	FSIM	Q score	PSNR	SSIM	FSIM	Q score
CAPE [17]	18.99	0.7435	0.8856	59.44	17.13	0.7853	0.8781	54.87
WVM [8]	17.70	0.8016	0.8695	53.17	11.25	0.5733	0.6072	41.12
L0S [46]	19.03	0.6644	0.7328	84.33	15.72	0.7311	0.7751	51.73
SMF [50]	18.61	0.7724	0.9035	81.07	14.85	0.6776	0.7622	50.77
DJF [22]	17.54	0.7395	0.9512	84.74	14.49	0.6736	0.7360	50.03
DRHT	28.18	0.9242	0.9622	97.87	22.60	0.7629	0.8691	56.17

Table 1: Quantitative evaluation on the standard datasets. The proposed DRHT method is compared with existing image correction methods based on several metrics including PSNR, SSIM, FSIM and Q score. It shows that the proposed DRHT method performs favorably against existing image correction methods.

Methods	City Scene dataset				Sun360 Outdoor dataset			
	PSNR	SSIM	FSIM	Q score	PSNR	SSIM	FSIM	Q score
Hdrcnn [6]	11.99	0.2249	0.5687	39.64	11.09	0.6007	0.8637	56.31
DrTMo [7]	-	-	-	-	14.64	0.6822	0.8101	52.39
DRHT	28.18	0.9242	0.9622	97.87	22.60	0.7629	0.8691	56.17

Table 2: Quantitative evaluation between the proposed DRHT method and two HDR prediction methods. The results of DrTMo on the City Scene Dataset are not available as it requires high resolution inputs. The evaluation indicates the proposed DRHT method is effective to generate HDR data compared with existing HDR prediction methods.

marks. The visual evaluation is shown in Figure 4 where the input images are captured in over exposure. The image filtering based methods are effective to preserve local edges. However, they cannot recover the details in the overexposed regions, as shown in (c), (d) and (f). It is because these methods tend to smooth the flat region while preserving the color contrast around the edge region. They fail to recover the details, which reside in the overexposed regions where the pixel values approach 255. Meanwhile, the image correction methods based on global contrast and saturation manipulation are not effective as shown in (r). They share the similar limitations as image filtering based methods as the pixel-level operation fails to handle overexposed images. The results of WVM [8] tend to be brighter as shown in (e), (m) and (u) as they over enhance the illumination layer decomposed from the input image. Compared with existing methods, the proposed DRHT method can successfully recover the missing details buried in the over exposed regions while maintaining the realistic global illumination.

Figure 5 shows some under/over exposed examples in the low-light scenes. It shows that the image filtering based methods can only strengthen existing details. CAPE [17] performs well in the low-light regions as shown in (b) but it simply adjusts the brightness and thus fails to correct all missing details. Figure 5(i) shows that WVM [8] performs poorly in the scenes with dark skies, as it fails to decompose the dark sky into reflectance and the illumination layers. Meanwhile, the missing details in the under/over exposed regions can be reconstructed via the proposed DRHT

method as shown in (h) and (p). Global illumination is also maintained through residual learning.

We note that the proposed DRHT method tends to slightly increase the intensity in the dark regions. There are two reasons for this. First, DRHT is trained on the city scene dataset [51], where the sun is always located near the center of the images. Hence, when the input image has some bright spots near to the center, the night sky will tend to appear brighter as shown in Figure 5(p)). Second, as we use the first network to predict the gamma compressed HDR image and then map it back to the LDR in the logarithmic domain, low intensity values may be increased through the inverse gamma mapping and logarithmic compression as shown in Figure 5(h).

In additional to visual evaluation, we also provide quantitative comparison between the proposed method and existing methods as summarized in Table 1. It shows that the proposed method performs favorably against existing methods under several numerical evaluation metrics.

We further compare the proposed DRHT method with two HDR prediction methods (*i.e.*, DrTMO [7] and Hdrcnn [6]). These two methods can be treated as image correction methods because their output HDR image can be tone mapped into the LDR image. In [7], two deep networks are proposed to first generate up-exposure and down-exposure LDR images from the single input LDR image. As each image with limited exposure cannot contain all the details of the scene to solve the under/over exposure problem, they fuse these multiple exposed images and use [18] to gen-

(a) Input   (b) DrTMo [7]   (c) Hdrcnn [6]   (d) DRHT   (e) Ground Truth

(e) Input   (f) DrTMo [7]   (g) Hdrcnn [6]   (h) DRHT   (e) Ground Truth
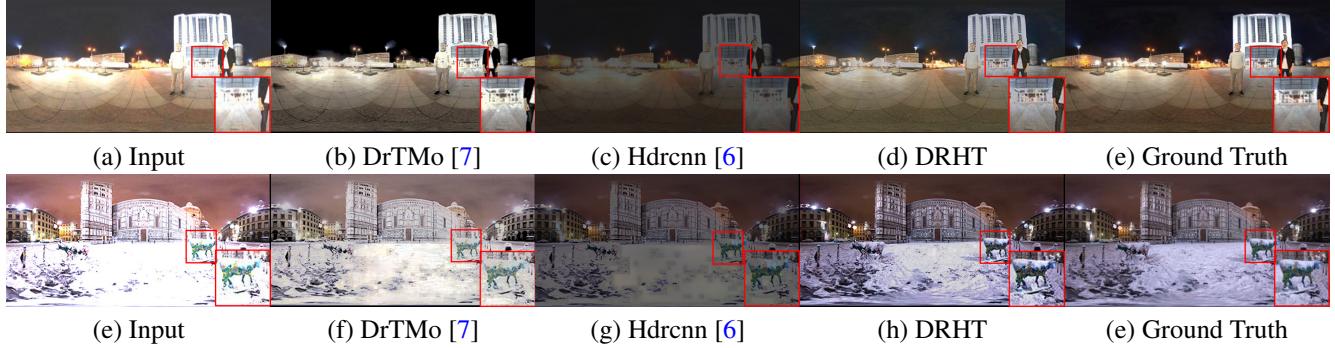
Figure 6: Visual comparison with two HDR based correction methods: DrTMo [7] and Hdrcnn [6], on the Sun360 outdoor dataset. The proposed DRHT performs better than these two methods in generating visually pleasing images.

erate the final LDR images. Eilertsen *et al*. [6] propose a deep network to blend the input LDR image with the reconstructed HDR information in order to recover the high dynamic range in the LDR output images. However, by using the highlight masks for blending, their method cannot deal with the under exposed regions and their results tend to be dim as shown in Figures 6(c) and 6(g). Meanwhile, we can also observe obvious flaws in the output images of both DrTMO [7] and Hdrcnn [6] (e.g., the man's white shirt in Figure 6(b) and the blocking effect in the snow in Figure 6(g)). The main reason lies in that existing tone mapping methods fail to preserve the local details from the HDR domain when the under/exposure exposure problem happens. In comparison, the proposed DRHT is effective to prevent this limitation because we do not attempt to recover the whole HDR image but only focus on recovering the missing details by residual learning. The quantitative evaluation results shown in Table 2 indicate that the proposed DRHT method performs favorably against these HDR prediction methods.

4.4. Limitation Analysis

Despite the aforementioned success, the proposed DRHT method contains limitation to recover the details when significant illumination contrast appears on the input images. Figure 7 shows one example. Although DRHT can effectively recover the missing details of the hut in the underexposed region (i.e., the red box in Figure 7), there are limited details around the sun (i.e., the black box). This is mainly because of the large area of overexposed sunshine is rare in our training dataset. In the future, we will augment our training dataset to incorporate such extreme cases to improve the performance.

5. Conclusion

In this paper, we propose a novel deep reciprocating HDR transformation (DRHT) model for under/over ex-

(a) Input   (b) DRHT
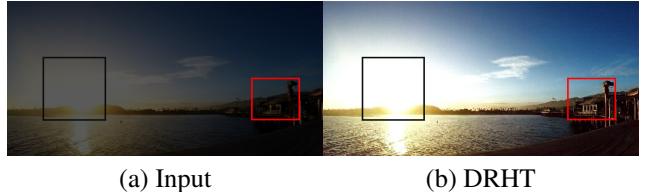
Figure 7: Limitation analysis. The proposed DRHT method is effective to recover the missing details in the underexposed region marked in the red box, while limits on the overexposed sunshine region marked in the black box.

posed image correction. We first trace back to the image formulation process to explain why the under/over exposure problem is observed in the LDR images, according to which we reformulate the image correction as the HDR mapping problem. We show that the buried details in the under/over exposed regions cannot be completely recovered in the LDR domain by existing image correction methods. Instead, the proposed DRHT method first revisits the HDR domain and recovers the missing details of natural scenes via the HDR estimation network, and then transfers the reconstructed HDR information back to the LDR domain to correct the image via another proposed LDR correction network. These two networks are formulated in an end-to-end manner as DRHT and achieve state-of-the-art correction performance on two benchmarks.

Acknowledgements

References

[1] L. Bao, Y. Song, Q. Yang, and N. Ahuja. An edge-preserving filtering framework for visibility restoration. In *International Conference on Pattern Recognition*, 2012.

[2] S. Bi, X. Han, and Y. Yu. An l_1 image transform for edge-preserving smoothing and scene-level intrinsic decomposition. *ACM Transactions on Graphics*, 2015.

[3] J. Cheng, Y.-H. Tsai, S. Wang, and M.-H. Yang. Segflow: Joint learning for video object segmentation and optical flow. In *IEEE International Conference on Computer Vision*, 2017.

[4] D.-A. Clevert, T. Unterthiner, and S. Hochreiter. Fast and accurate deep network learning by exponential linear units (elus). *arXiv:1511.07289*, 2015.

[5] P. Debevec and J. Malik. Recovering high dynamic range radiance maps from photographs. In *ACM Transactions on Graphics (SIGGRAPH)*, 2008.

[6] G. Eilertsen, J. Kronander, G. Denes, R. Mantiuk, and J. Unger. Hdr image reconstruction from a single exposure using deep cnns. *ACM Transactions on Graphics*, 2017.

[7] Y. Endo, Y. Kanamori, and J. Mitani. Deep reverse tone mapping. *ACM Transactions on Graphics*, 2017.

[8] X. Fu, D. Zeng, Y. Huang, X.-P. Zhang, and X. Ding. A weighted variational model for simultaneous reflectance and illumination estimation. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2016.

[9] Google. Tensorflow.

[10] X. Guo, Y. Li, and H. Ling. Lime: Low-light image enhancement via illumination map estimation. *IEEE Transactions on Image Processing*, 2017.

[11] K. He, J. Sun, and X. Tang. Guided image filtering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013.

[12] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2016.

[13] S. He, J. Jiao, X. Zhang, G. Han, and R. Lau. Delving into salient object subitizing and detection. In *IEEE International Conference on Computer Vision*, 2017.

[14] S. J. Hwang, A. Kapoor, and S. B. Kang. Context-based automatic local image enhancement. In *European Conference on Computer Vision*, 2012.

[15] E. Ilg, N. Mayer, T. Saikia, M. Keuper, A. Dosovitskiy, and T. Brox. Flownet 2.0: Evolution of optical flow estimation with deep networks. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2017.

[16] S. Ioffe and C. Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International Conference on Machine Learning*, 2015.

[17] L. Kaufman, D. Lischinski, and M. Werman. Content-aware automatic photo enhancement. *Computer Graphics Forum*, 2012.

[18] M. Kim and J. Kautz. Consistent tone reproduction. In *International Conference on Computer Graphics and Imaging*, 2008.

[19] Y. Kim, H. Jung, D. Min, and K. Sohn. Deeply aggregated alternating minimization for image restoration. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2017.

[20] P. Kingma and J. Ba. Adam: A Method for Stochastic Optimization. *arXiv:1412.6980*, 2014.

[21] J. Li, X. Chen, D. Zou, B. Gao, and W. Teng. Conformal and low-rank sparse representation for image restoration. In *IEEE International Conference on Computer Vision*, 2015.

[22] Y. Li, J.-B. Huang, N. Ahuja, and M.-H. Yang. Deep joint image filtering. In *European Conference on Computer Vision*, 2016.

[23] S. Liu, J. Pan, and M.-H. Yang. Learning recursive filters for low-level vision via a hybrid neural network. In *European Conference on Computer Vision*, 2016.

[24] Z. Ma, K. He, Y. Wei, J. Sun, and E. Wu. Constant time weighted median filtering for stereo matching and beyond. In *IEEE International Conference on Computer Vision*, 2013.

[25] R. Mantiuk, K. Joong, A. Rempel, and W. Heidrich. Hdr-vdp-2: A calibrated visual metric for visibility and quality predictions in all luminance conditions. *ACM Transactions on Graphics*, 2011.

[26] X. Mei, H. Qi, B.-G. Hu, and S. Lyu. Improving image restoration with soft-rounding. In *IEEE International Conference on Computer Vision*, 2015.

[27] M. Pharr, W. Jakob, and G. Humphreys. *Physically based rendering: From theory to implementation*. 2016.

[28] S. Pizer, E. Amburn, J. Austin, R. Cromartie, A. Geselowitz, T. Greer, H. Bart, J. Zimmerman, and K. Zuiderveld. Adaptive histogram equalization and its variations. *Computer Vision, Graphics, and Image Processing*, 1987.

[29] W. Ren, S. Liu, H. Zhang, J. Pan, X. Cao, and M.-H. Yang. Single image dehazing via multi-scale convolutional neural networks. In *European Conference on Computer Vision*, 2016.

[30] W. Ren, J. Pan, X. Cao, and M.-H. Yang. Video deblurring via semantic segmentation and pixel-wise non-linear kernel. In *IEEE International Conference on Computer Vision*, 2017.

[31] A. Rivera, B. Ryu, and O. Chae. Content-aware dark image enhancement through channel division. *IEEE Transactions on Image Processing*, 2012.

[32] L. Shen, Z. Yue, F. Feng, Q. Chen, S. Liu, and J. Ma. MSR-net:Low-light Image Enhancement Using Deep Convolutional Network. *arXiv:1711.02488*, 2017.

[33] X. Shen, C. Zhou, L. Xu, and J. Jia. Mutual-structure for joint filtering. In *IEEE International Conference on Computer Vision*, 2015.

[34] M. Son, Y. Lee, H. Kang, and S. Lee. Art-photographic detail enhancement. *Computer Graphics Forum*, 2014.

[35] Y. Song, L. Bao, S. He, Q. Yang, and M.-H. Yang. Stylizing face images via multiple exemplars. *Computer Vision and Image Understanding*, 2017.

[36] Y. Song, L. Bao, X. Xu, and Q. Yang. Decolorization: Is rgb2gray () out? In *ACM SIGGRAPH Asia Technical Briefs*, 2013.

[37] Y. Song, L. Bao, and Q. Yang. Real-time video decolorization using bilateral filtering. In *IEEE Winter Conference on Applications of Computer Vision*, 2014.

[38] Y. Song, C. Ma, L. Gong, J. Zhang, R. Lau, and M.-H. Yang. Crest: Convolutional residual learning for visual tracking. In *IEEE International Conference on Computer Vision*, 2017.

[39] Y. Song, J. Zhang, L. Bao, and Q. Yang. Fast preprocessing for robust face sketch synthesis. In *International Joint Conference on Artificial Intelligence*, 2017.

[40] Y. Tai, J. Yang, X. Liu, and C. Xu. Memnet: A persistent memory network for image restoration. In *IEEE International Conference on Computer Vision*, 2017.

[41] Y.-H. Tsai, X. Shen, Z. Lin, K. Sunkavalli, X. Lu, and M.-H. Yang. Deep image harmonization. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2017.

[42] B. Wang, Y. Yu, T.-T. Wong, C. Chen, and Y.-Q. Xu. Data-driven image color theme enhancement. *ACM Transactions on Graphics*, 2010.

[43] B. Wang, Y. Yu, and Y.-Q. Xu. Example-based image color and tone style enhancement. *ACM Transactions on Graphics*, 2011.

[44] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 2004.

[45] J. Xiao, K. Ehinger, A. Oliva, and A. Torralba. Recognizing scene viewpoint using panoramic place representation. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2012.

[46] L. Xu, C. Lu, Y. Xu, and J. Jia. Image smoothing via l0 gradient minimization. *ACM Transactions on Graphics*, 2011.

[47] L. Xu, Q. Yan, Y. Xia, and J. Jia. Structure extraction from texture via relative total variation. *ACM Transactions on Graphics*, 2012.

[48] J. Yan, S. Lin, S. Bing Kang, and X. Tang. A learning-to-rank approach for image color enhancement. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2014.

[49] Z. Yan, H. Zhang, B. Wang, S. Paris, and Y. Yu. Automatic photo adjustment using deep neural networks. *ACM Transactions on Graphics*, 2015.

[50] Q. Yang. Semantic filtering. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2016.

[51] J. Zhang and J.-F. Lalonde. Learning high dynamic range from outdoor panoramas. In *IEEE International Conference on Computer Vision*, 2017.

[52] K. Zhang, W. Zuo, S. Gu, and L. Zhang. Learning deep cnn denoiser prior for image restoration. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2017.

[53] L. Zhang, L. Zhang, X. Mou, and D. Zhang. Fsim: A feature similarity index for image quality assessment. *IEEE Transactions on Image Processing*, 2011.

[54] Q. Zhang, X. Shen, L. Xu, and J. Jia. Rolling guidance filter. In *European Conference on Computer Vision*, 2014.

[55] Q. Zhang, L. Xu, and J. Jia. 100+ times faster weighted median filter (wmf). In *IEEE Conference on Computer Vision and Pattern Recognition*, 2014.