GHOST-FREE HDR IMAGING VIA UNROLLING LOW-RANK MATRIX COMPLETION

Truong Thanh Nhat Mai[†], *Edmund Y. Lam*[‡], *and Chul Lee*[†]

[†]Department of Multimedia Engineering, Dongguk University, Seoul, Korea
[‡]Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong

ABSTRACT

We propose a ghost-free high dynamic range (HDR) image synthesis algorithm by unrolling low-rank matrix completion. By exploiting the low-rank structure of the irradiance maps from low dynamic range (LDR) images, we formulate ghost-free HDR imaging as a general low-rank matrix completion problem. Then, we solve the problem iteratively using the augmented Lagrange multiplier (ALM) method. At each iteration, the optimization variables are updated by closed-form solutions and the regularizers are updated by learned deep neural networks. Experimental results show that the proposed algorithm provides better image qualities with fewer visual artifacts compared to state-of-the-art algorithms.

***Index Terms*—** High dynamic range imaging, unrolled optimization, low-rank matrix completion.

1. INTRODUCTION

The advancements in digital imaging technology have enabled the capture of high-quality images. However, because the dynamic range of natural scenes often exceeds those of the sensors [1], captured images contain under- and over-exposed regions, degrading image quality. A common approach to overcoming the limited dynamic range of the sensors is to synthesize a high dynamic range (HDR) image by merging multiple low dynamic range (LDR) images taken with different exposure times. However, camera or object motions between different exposures cause ghosting artifacts, which degrade the quality of the synthesized images. Therefore, due to its practical importance, a lot of research has been carried out to synthesize high-quality HDR images without ghosting artifacts, which is called ghost-free HDR imaging.

Recent ghost-free HDR imaging algorithms can be broadly categorized into model-based and learning-based approaches. Model-based approaches are based on ghosting artifacts generation models and have developed algorithms to remove the artifacts. For example, in [2, 3], ghost region detection was employed to alleviate contributions of the region to the synthesized HDR image. In [4, 5], input LDR images are first aligned and then are merged to provide an HDR image. Based on the assumption of a static underlying scene, rank minimization has been employed for HDR synthesis [6, 7]. Despite their theoretical completeness, rank minimization-based approaches may fail if the number of input LDR images is insufficient. A common disadvantage of these approaches is that the synthesis performance becomes degraded if the models inaccurately fit real-world scenarios.

Inspired by the recent success of deep learning in various computer vision and image processing tasks, several deep learning-based HDR image synthesis algorithms have been developed [8–10]. These algorithms employ convolutional neural networks (CNNs) that take a set of multiple LDR images and produce an HDR image in an end-to-end manner. Such approaches can overcome the limitations of hand-crafted models through the learned high-dimensional features from training data. However, the performance of learning-based approaches are affected by the quality and diversity of training data [11]. Moreover, due to their black box nature, it is usually difficult to interpret the behaviors of CNNs.

Recently, a novel technique called algorithm unrolling [12] was developed to connect model-based iterative algorithms to CNN architectures, overcoming their weaknesses while taking strengths. Specifically, an iterative algorithm is expanded into layers of a deep network, in which each layer represents an iteration of the algorithm. As data flow through CNN is analogous to running the iterative algorithm, the learning-based algorithms inherit strong theoretical bases from the model-based algorithms. Because of its interpretable CNN architectures from reasonably sized training sets, algorithm unrolling has been applied in various image processing tasks [13–15].

In this work, we propose a ghost-free HDR image synthesis algorithm by unrolling low-rank matrix completion that takes advantage of benefits from both model-based and learning-based approaches. First, we formulate ghost-free HDR imaging as a general matrix completion problem, assuming linear dependency among LDR images. Then, we solve the problem using the augmented Lagrange multiplier (ALM) method by reformulating it into a series of subproblems and iteratively solving them. At each iteration, the optimization variables are updated by closed-form solutions, while the regularizes are updated by learned CNNs. Experimental results show that the proposed algorithm achieves higher-quality HDR images with fewer visual artifacts compared to state-of-the-art algorithms [7–10].

The remainder of this paper is organized as follows: Section 2 describes the proposed HDR image synthesis algorithm, and Section 3 discusses experimental results. Finally, Section 4 concludes this work.

2. PROPOSED ALGORITHM

In this section, we formulate HDR image synthesis as a rank minimization problem and then solve it using an unrolled optimization.

2.1. Problem Formulation

Given a set of images taken with different exposure times, we synthesize an HDR image while suppressing ghosting artifacts caused by moving objects. First, the images are warped to the reference image, which is at the middle of the sequence, using SIFT-Flow [16], producing a set of warped images $\mathbf{I} = \{\mathrm{vec}(\mathbf{I}_1), \ldots, \mathrm{vec}(\mathbf{I}_n)\}$, where $\mathrm{vec}(\mathbf{I}_i) \in \mathbb{R}^m$ denotes a vector of pixel values and n is the number of images. Then, we construct the observed irradiance ma-

trix $\mathbf{D} = [\mathrm{vec}(\mathbf{H}_1), \ldots, \mathrm{vec}(\mathbf{H}_n)]$, where $\mathrm{vec}(\mathbf{H}_i)$ denotes the irradiance vector for the ith warped image.

The HDR synthesis then can be formulated as a low-rank matrix completion [7]. Specifically, based on the assumption that the scene irradiance matrix $\mathbf{D}$ can be decomposed into the background and foreground components $\mathbf{X}$ and $\mathbf{E}$, respectively, and that the background is static and the foreground object is small, $\mathbf{X}$ has a low rank and $\mathbf{E}$ is a sparse matrix. Further, due to under- and over-exposed regions in input images, only a limited number of pixels in a set $\mathbf{\Omega}$ is observed. Therefore, using the truncated nuclear norm $\|\mathbf{X}\|_r = \sum_{i=r+1}^{\min(m,n)} \sigma_i(\mathbf{X})$, where $\sigma_i(\mathbf{X})$ is the ith largest singular value of $\mathbf{X}$, as an approximation for the matrix rank [17], the estimation of the low-rank scene background matrix for ghost-free HDR image can be approximately formulated as [7]

$$\begin{aligned} \underset{\mathbf{X},\mathbf{E}}{\text{minimize}} \quad & \|\mathbf{X}\|_r + \lambda\|\mathbf{E}\|_1 \\ \text{subject to} \quad & \mathcal{P}_{\mathbf{\Omega}}(\mathbf{X}+\mathbf{E}) = \mathcal{P}_{\mathbf{\Omega}}(\mathbf{D}), \end{aligned} \tag{1}$$

where λ controls the relative importance between the two terms, and $\mathcal{P}_{\mathbf{\Omega}}$ denotes a sampling operator in the observed region $\mathbf{\Omega}$ as

$$[\mathcal{P}_{\mathbf{\Omega}}(\mathbf{X})]_{ij} = \begin{cases} \mathbf{X}_{ij}, & \text{if } (i,j) \in \mathbf{\Omega}, \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

Based on the assumption that the background scene is static, we fix r in (1) to 1 in this work.

Note that, in (1), the models for $\mathbf{X}$ and $\mathbf{E}$, *i.e.*, low-rankness of the scene and sparse error due to inaccurate warping, respectively, are made based on the assumptions. However, these hand-crafted priors may fail to accurately capture all features of the scene and errors, such as color distribution or underlying structures. To overcome the limitation of the models, we add two general regularization terms $f(\cdot)$ and $g(\cdot)$ for $\mathbf{X}$ and $\mathbf{E}$, respectively, which are learned from data. We also consider noise in the observed pixels for real-world acquisition as done in [7], *i.e.*, $\mathcal{P}_{\mathbf{\Omega}}(\mathbf{X}+\mathbf{E}+\mathbf{N}) = \mathcal{P}_{\mathbf{\Omega}}(\mathbf{D})$, where $\mathbf{N}$ is a noise matrix. Then, by introducing slack variables, the optimization problem in (1) is rewritten as

$$\begin{aligned} \underset{\mathbf{X},\mathbf{E},\mathbf{S}}{\text{minimize}} \quad & \|\mathbf{X}\|_r + \lambda\|\mathbf{E}\|_1 + f(\mathbf{X}) + g(\mathbf{E}) \\ \text{subject to} \quad & \mathbf{X}+\mathbf{E}+\mathbf{S} = \mathcal{P}_{\mathbf{\Omega}}(\mathbf{D}), \\ & \|\mathcal{P}_{\mathbf{\Omega}}(\mathbf{S})\|_F \le \delta, \end{aligned} \tag{3}$$

where $\mathbf{S}$ is a matrix of slack variable, and $\delta \ge 0$ is the noise level.

In (3), we define the observed region $\mathbf{\Omega}$ as a set of reliable pixel locations in terms of pixel value and structure information. First, we define the set of well-exposed pixel locations $\mathbf{\Omega}_e$ as

$$\mathbf{\Omega}_e = \{(i,j) \mid 0.10 \le \mathbf{I}(i,j) \le 0.99\}. \tag{4}$$

Next, we employ the structure similarity index (SSIM) [18] to consider the structural information, as it evaluates the degree of similarity between two images. Specifically, we define the set of well-warped pixel locations $\mathbf{\Omega}_w$ based on the SSIM score computed between the reference and warped images, given by

$$\mathbf{\Omega}_w = \{(i,j) \mid \mathrm{SSIM}(\mathbf{I}(i,j)) \ge 0.90\}. \tag{5}$$

An SSIM score is computed for each pixel using an 11×11 window. Then, we define the observed region as $\mathbf{\Omega} = \mathbf{\Omega}_e \cap \mathbf{\Omega}_w$.

2.2. Iterative Solutions

We employ the ALM method to solve the optimization in (3). To this end, we first reformulate it to split the optimization variables into two parts by introducing the auxiliary variables $\mathbf{P}$ and $\mathbf{Q}$ as

$$\begin{aligned} \underset{\mathbf{X},\mathbf{E},\mathbf{S},\mathbf{P},\mathbf{Q}}{\text{minimize}} \quad & \|\mathbf{X}\|_r + \lambda\|\mathbf{E}\|_1 + f(\mathbf{P}) + g(\mathbf{Q}) \\ \text{subject to} \quad & \mathbf{P} = \mathbf{X}, \mathbf{Q} = \mathbf{E}, \\ & \mathbf{X}+\mathbf{E}+\mathbf{S} = \mathcal{P}_{\mathbf{\Omega}}(\mathbf{D}), \\ & \|\mathcal{P}_{\mathbf{\Omega}}(\mathbf{S})\|_F \le \delta. \end{aligned} \tag{6}$$

Then, we define the augmented Lagrangian function $\mathcal{L}$ for (6) as

$$\begin{aligned} \mathcal{L}(\mathbf{X},\mathbf{E},\mathbf{S},\mathbf{P},\mathbf{Q},\mathbf{\Lambda},\mathbf{\Gamma},\mathbf{\Phi}) &= \|\mathbf{X}\|_r + \lambda\|\mathbf{E}\|_1 + f(\mathbf{P}) + g(\mathbf{Q}) \\ &+ \langle \mathbf{\Lambda}, \mathcal{P}_{\mathbf{\Omega}}(\mathbf{D}) - \mathbf{X} - \mathbf{E} - \mathbf{S} \rangle + \frac{\mu}{2}\|\mathcal{P}_{\mathbf{\Omega}}(\mathbf{D}) - \mathbf{X} - \mathbf{E} - \mathbf{S}\|_F^2 \\ &+ \langle \mathbf{\Gamma}, \mathbf{X} - \mathbf{P} \rangle + \frac{\alpha}{2}\|\mathbf{X} - \mathbf{P}\|_F^2 + \langle \mathbf{\Phi}, \mathbf{E} - \mathbf{Q} \rangle + \frac{\beta}{2}\|\mathbf{E} - \mathbf{Q}\|_F^2, \end{aligned} \tag{7}$$

where μ, α, and $\beta > 0$ are penalty parameters, $\mathbf{\Lambda}$, $\mathbf{\Gamma}$, and $\mathbf{\Phi} \in \mathbb{R}^{m \times n}$ are Lagrange multiplier matrices, and $\langle \mathbf{A}, \mathbf{B} \rangle = \mathrm{tr}(\mathbf{A}\mathbf{B}^T)$ denotes the matrix inner product.

Solutions to the optimization problem in (6) can be obtained by minimizing the augmented Lagrangian function $\mathcal{L}$ in (7). More specifically, we employ the alternating direction method of multipliers [19], which split the optimization over variables $\mathbf{X}$, $\mathbf{E}$, $\mathbf{S}$, $\mathbf{P}$, and $\mathbf{Q}$ and multipliers $\mathbf{\Lambda}$, $\mathbf{\Gamma}$, and $\mathbf{\Phi}$ and then solve them one-by-one iteratively. We describe how each subproblem is solved.

X-subproblem: In the first step, we update $\mathbf{X}$ as

$$\begin{aligned} \mathbf{X}_{k+1} &= \arg\min_{\mathbf{X}} \mathcal{L}(\mathbf{X}, \mathbf{E}_k, \mathbf{S}_k, \mathbf{P}_k, \mathbf{Q}_k, \mathbf{\Lambda}_k, \mathbf{\Gamma}_k, \mathbf{\Phi}_k) \\ &= \arg\min_{\mathbf{X}} \|\mathbf{X}\|_r + \frac{\sqrt{\mu_k + \alpha_k}}{2}\|\mathbf{X} - \mathbf{\Psi}_{\mathbf{X},k}\|_F^2, \end{aligned} \tag{8}$$

where $\mathbf{\Psi}_{\mathbf{X},k} = (\mu_k + \alpha_k)^{-1}(\mathbf{\Lambda}_k + \mu_k\mathcal{P}_{\mathbf{\Omega}}(\mathbf{D}) - \mu_k\mathbf{E}_k - \mu_k\mathbf{S}_k + \alpha_k\mathbf{P}_k - \mathbf{\Gamma}_k)$. We employ the partial singular value thresholding (PSVT) operator [6] to solve the problem in (8). Specifically, given the singular value decomposition of a matrix $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$, where $\mathbf{\Sigma} = \mathrm{diag}(\sigma_1, \ldots, \sigma_{\min(m,n)})$, the PSVT operator $\mathbb{P}_{r,\tau}$ is defined as $\mathbb{P}_{r,\tau}(\mathbf{A}) = \mathbf{U}(\mathbf{\Sigma}_1 + \mathcal{S}_\tau(\mathbf{\Sigma}_2))\mathbf{V}^T$, where $\mathbf{\Sigma}_1 = \mathrm{diag}(\sigma_1, \ldots, \sigma_r, 0, \ldots, 0)$, $\mathbf{\Sigma}_2 = \mathrm{diag}(0, \ldots, 0, \sigma_{r+1}, \ldots, \sigma_{\min(m,n)})$, and $\mathcal{S}_\tau(\cdot)$ denotes the element-wise soft-thresholding operator [20] for $\tau > 0$; that is, $[\mathcal{S}(\mathbf{A})]_{ij} = \mathrm{sign}(A_{ij}) \cdot \max\{|A_{ij}| - \tau, 0\}$. Then, the closed-form solution to (8) is given by [6]

$$\mathbf{X}_{k+1} = \mathbb{P}_{r, \frac{1}{\sqrt{\mu_k + \alpha_k}}}(\mathbf{\Psi}_{\mathbf{X},k}). \tag{9}$$

E-subproblem: Next, we estimate $\mathbf{E}$ by solving

$$\begin{aligned} \mathbf{E}_{k+1} &= \arg\min_{\mathbf{E}} \mathcal{L}(\mathbf{X}_{k+1}, \mathbf{E}, \mathbf{S}_k, \mathbf{P}_k, \mathbf{Q}_k, \mathbf{\Lambda}_k, \mathbf{\Gamma}_k, \mathbf{\Phi}_k) \\ &= \arg\min_{\mathbf{E}} \lambda\|\mathbf{E}\|_1 + \frac{\sqrt{\mu_k + \beta_k}}{2}\|\mathbf{E} - \mathbf{\Psi}_{\mathbf{E},k}\|_F^2, \end{aligned} \tag{10}$$

where $\mathbf{\Psi}_{\mathbf{E},k} = (\mu_k + \beta_k)^{-1}(\mathbf{\Lambda}_k + \mu_k\mathcal{P}_{\mathbf{\Omega}}(\mathbf{D}) - \mu_k\mathbf{X}_{k+1} - \mu_k\mathbf{S}_k + \beta_k\mathbf{Q}_k - \mathbf{\Phi}_k)$. The closed-form solution to (10) can be obtained by the soft-thresholding operator [20], given by

$$\mathbf{E}_{k+1} = \mathcal{S}_{\frac{\lambda}{\sqrt{\mu_k + \beta_k}}}(\mathbf{\Psi}_{\mathbf{E},k}). \tag{11}$$

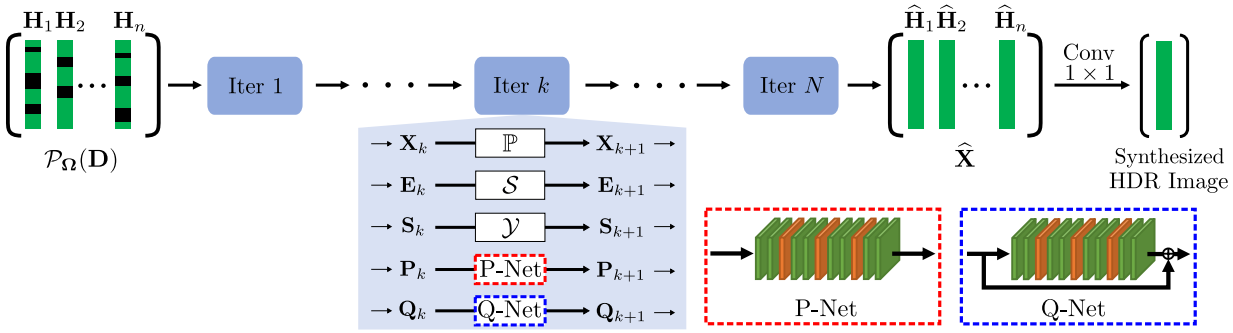

Fig. 1. Overview of the proposed unrolled optimization algorithm for HDR image synthesis. The green and orange blocks in P-Net and Q-Net indicate the convolutional layers and ReLU layers, respectively.

S-subproblem: We also update $\mathbf{S}$ by solving the following optimization problem.

$$\begin{aligned}\mathbf{S}_{k+1} &= \underset{\|\mathcal{P}_{\mathbf{\Omega}}(\mathbf{S})\|_F \le \delta}{\arg\min}\ \mathcal{L}(\mathbf{X}_{k+1}, \mathbf{E}_{k+1}, \mathbf{S}, \mathbf{P}_k, \mathbf{Q}_k, \mathbf{\Lambda}_k, \mathbf{\Gamma}_k, \mathbf{\Phi}_k) \\ &= \underset{\|\mathcal{P}_{\mathbf{\Omega}}(\mathbf{S})\|_F \le \delta}{\arg\min}\ \|\mathbf{S} - \mathbf{Y}_k\|_F^2, \end{aligned} \tag{12}$$

where $\mathbf{Y}_k = \mathcal{P}_{\mathbf{\Omega}}(\mathbf{D}) - \mathbf{X}_{k+1} - \mathbf{E}_{k+1} + \mu_k^{-1}\mathbf{\Lambda}_k$. It was shown in [7] that the closed-form solution to (12) is given by

$$\mathbf{S}_{k+1} = \mathcal{Y}(\mathbf{Y}_k) = \mathcal{P}_{\mathbf{\Omega}^c}(\mathbf{Y}_k) + \min\left\{\frac{\delta}{\|\mathcal{P}_{\mathbf{\Omega}}(\mathbf{Y}_k)\|_F}, 1\right\}\mathcal{P}_{\mathbf{\Omega}}(\mathbf{Y}_k). \tag{13}$$

P- and Q-subproblems: Then, $\mathbf{P}$ and $\mathbf{Q}$ are updated as

$$\begin{aligned}\mathbf{P}_{k+1} &= \arg\min_{\mathbf{P}}\ \mathcal{L}(\mathbf{X}_{k+1}, \mathbf{E}_{k+1}, \mathbf{S}_{k+1}, \mathbf{P}, \mathbf{Q}_k, \mathbf{\Lambda}_k, \mathbf{\Gamma}_k, \mathbf{\Phi}_k) \\ &= \arg\min_{\mathbf{P}}\ f(\mathbf{P}) + \frac{\sqrt{\alpha_k}}{2}\|\mathbf{P} - (\mathbf{X}_{k+1} + \alpha_k^{-1}\mathbf{\Gamma}_k)\|_F^2, \end{aligned} \tag{14}$$

$$\begin{aligned}\mathbf{Q}_{k+1} &= \arg\min_{\mathbf{Q}}\ \mathcal{L}(\mathbf{X}_{k+1}, \mathbf{E}_{k+1}, \mathbf{S}_{k+1}, \mathbf{P}_{k+1}, \mathbf{Q}, \mathbf{\Lambda}_k, \mathbf{\Gamma}_k, \mathbf{\Phi}_k) \\ &= \arg\min_{\mathbf{Q}}\ g(\mathbf{Q}) + \frac{\sqrt{\beta_k}}{2}\|\mathbf{Q} - (\mathbf{E}_{k+1} + \beta_k^{-1}\mathbf{\Phi}_k)\|_F^2. \end{aligned} \tag{15}$$

Because the functions $f(\cdot)$ and $g(\cdot)$ are learned from data and thus nondeterministic, we develop CNNs that obtain the solutions to the optimization problems in (14) and (15). We denote the CNNs at the kth iteration as P-Net$_k$ and Q-Net$_k$. Then, we represent the closed-form solutions to the optimization problem in (14) and (15) as

$$\mathbf{P}_{k+1} = \text{P-Net}_k\left(\mathbf{X}_{k+1} + \alpha_k^{-1}\mathbf{\Gamma}_k\right), \tag{16}$$

$$\mathbf{Q}_{k+1} = \text{Q-Net}_k\left(\mathbf{E}_{k+1} + \beta_k^{-1}\mathbf{\Phi}_k\right). \tag{17}$$

P-Net$_k$ and Q-Net$_k$ are equivalent to the proximal operators corresponding to the regularization terms. During training, the weights of P-Net$_k$ and Q-Net$_k$ are adjusted accordingly to their input matrices $\mathbf{X}_{k+1} + \alpha_k^{-1}\mathbf{\Gamma}_k$ and $\mathbf{E}_{k+1} + \beta_k^{-1}\mathbf{\Phi}_k$, respectively, to produce the optimal solutions $\mathbf{P}_{k+1}$ and $\mathbf{Q}_{k+1}$. The details of P-Net and Q-Net are described in Section 2.3.

Finally, the Lagrange multiplier matrices are updated as

$$\mathbf{\Lambda}_{k+1} = \mathbf{\Lambda}_k + \mu_k(\mathcal{P}_{\mathbf{\Omega}}(\mathbf{D}) - \mathbf{X}_{k+1} - \mathbf{E}_{k+1} - \mathbf{S}_{k+1}), \tag{18}$$

$$\mathbf{\Gamma}_{k+1} = \mathbf{\Gamma}_k + \alpha_k(\mathbf{X}_{k+1} - \mathbf{P}_{k+1}), \tag{19}$$

$$\mathbf{\Phi}_{k+1} = \mathbf{\Phi}_k + \beta_k(\mathbf{E}_{k+1} - \mathbf{Q}_{k+1}). \tag{20}$$

2.3. Implementation

Fig. 1 shows an overview of the proposed unrolled optimization algorithm for HDR image synthesis. An observed matrix, $\mathbf{X}_1 = \mathcal{P}_{\mathbf{\Omega}}(\mathbf{D})$, is fed as input and all the other optimization variables are initialized to zeros. The iterative algorithm is unrolled into N blocks, and each block updates the optimization variables for subproblems in the previous section at the corresponding iteration. The resulting HDR image is synthesized by applying a 1×1 convolutional layer to the columns of the estimated background irradiance maps $\widehat{\mathbf{X}}$.

Note that $\mathbf{P}$ is a regularization variable for the final result $\mathbf{X}$. Therefore, we use a series of convolutional layers and the rectified linear unit (ReLU) layers for extensive feature extraction for P-Net. Each convolutional layer has 64 filters of size 3×3. On the other hand, $\mathbf{Q}$ regularizes the structure of the error component $\mathbf{E}$, which can be regarded as residual information. Thus, we employ a residual CNN for Q-Net. Specifically, the structure of Q-Net is identical to that of P-Net except for a skip connection. Fig. 1 shows the detailed structures of P-Net and Q-Net.

To train P-Net and Q-Net, we compute the synthesis loss L_{HDR} as the L_1-norm of the difference between a ground-truth HDR image $\mathbf{H}_{\text{gt}}$ and a synthesized HDR image $\widehat{\mathbf{H}}$, given by

$$L_{\text{HDR}} = \|l(\widehat{\mathbf{H}}) - l(\mathbf{H}_{\text{gt}})\|_1, \tag{21}$$

where $l(\cdot)$ denotes the irradiance value conversion to the perceptually uniform domain [21].

3. EXPERIMENTAL RESULTS

We evaluate the performance of the proposed HDR image synthesis algorithm on the HDR-HDM dataset [22]. We randomly chose three consecutive frames in the videos, then generated multi-exposure images for each set using exposure biases $\{-3, 0, +3\}$. We finally obtained 187 multi-exposure image sets with ground-truth HDR images as the middle images. We then randomly selected 132 sets for training and 55 sets for testing.

We compare the performance of the proposed algorithm with those of TNNM-ALM [7], Kalantari and Ramamoorthi's algorithm [8], Wu *et al.*'s algorithm [9], and AHDRNet [10]. While TNNM-ALM is a rank minimization model-based algorithm, the others are deep learning-based algorithms. The source codes for these algorithms are provided by the respective authors, and they were retrained with the new dataset. We used the tone-mapping technique [23] to print the synthesized HDR images. The number of

(a) (b) (c) (d) (e) (f)

Fig. 2. Comparison of the HDR synthesis results. (a) Ground-truth, and results of (b) TNNM-ALM [7], (c) Kalantari and Ramamoorthi's algorithm [8], (d) Wu *et al.*'s algorithm [9], (e) AHDRNet [10], and (f) the proposed algorithm. The second and fourth rows show the magnified parts for the red rectangles in the first and third rows, respectively.

Table 1. Quantitative comparison of the HDR synthesis performance using three objective quality metrics: PU-PSNR, PU-MSSSIM, and HDR-VDP.

	PU-PSNR	PU-MSSSIM	HDR-VDP
TNNM-ALM [7]	34.84	0.9640	66.32
Kalantari [8]	37.96	0.9894	66.55
Wu *et al.* [9]	43.11	0.9950	73.50
AHDRNet [10]	41.08	0.9886	71.78
Proposed	48.80	0.9983	74.00

unrolled iterations N is set to 10. The parameters λ and δ in (6) and μ_k, α_k, and β_k in (7) are all learned from training data.

Fig. 2 compares the synthesis results and their detailed parts on the test images. TNNM-ALM [7] in Fig. 2(b) yields strong color artifacts. This is because, while rank minimization requires large number of columns in the data matrix to provide high-quality results, only three columns (exposures) are used in this work. In Fig. 2(c), Kalantari and Ramamoorthi's algorithm [8] fails to reconstruct the image details in the saturated regions because it learns the weights for averaging warped radiance values. Wu *et al.*'s algorithm [9] and AHDRNet [10] in Figs. 2(d) and (e), respectively, provide better results. However, the results lose some textures and show color differences from ground-truth, since end-to-end learning infers them from learned features, which may be different from those in the ground-truth. Furthermore, AHDRNet yields ghosting artifacts in the blue rectangle in the fourth row due to incorrect estimation of the attention map. In contrast, the proposed algorithm in Fig. 2(f) faithfully synthesizes HDR images without texture losses and color artifacts by strictly constraining low-rank priors with learned regularizers.

We also perform objective evaluation of the proposed algorithm in comparison with the conventional algorithms using three quality metrics: perceptually uniform extension to PSNR (PU-PSNR) and multi-scale structural similarity index (PU-MSSSIM) [24], and high dynamic range visible difference predictor (HDR-VDP) [25]. Table 1 quantitatively compares the synthesis performance of the proposed algorithm against those of the conventional algorithms for the test images. For each metric, a higher score indicates better results. As the proposed algorithm effectively reconstructs missing information using low-rank priors and learned regularizers, it provides significantly higher PU-PSNR and PU-MSSSIM scores than the conventional algorithms. Also, the proposed algorithm achieves the best performance in terms of the HDR-VDP scores, providing the least perceptual differences.

4. CONCLUSIONS

We developed a ghost-free HDR imaging algorithm by unrolling low-rank matrix completion to take advantage of both model-based and learning-based approaches. We first formulated a general low-rank matrix completion problem with advanced regularizers, and then solved it efficiently by employing an iterative algorithm. Specifically, the proposed algorithm comprises a series of blocks, corresponding to each iteration, where the optimization variables are updated by closed-form solutions and the regularizers are updated by learned CNNs. Experimental results demonstrated that the proposed algorithm achieves higher HDR image synthesis performance than state-of-the-art algorithms.

5. REFERENCES

[1] C. A. Metzler, H. Ikoma, Y. Peng, and G. Wetzstein, "Deep optics for single-shot high-dynamic-range imaging," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2020, pp. 1372–1382.

[2] Y. S. Heo, K. M. Lee, S. U. Lee, Y. Moon, and J. Cha, "Ghost-free high dynamic range imaging," in *Proc. Asian Conf. Comput. Vis.*, Nov. 2010, pp. 486–500.

[3] S. Raman and S. Chaudhuri, "Reconstruction of high contrast images for dynamic scenes," *Vis. Comput.*, vol. 27, no. 12, pp. 1099–1114, Dec. 2011.

[4] D. Hafner, O. Demetz, and J. Weickert, "Simultaneous HDR and optic flow computation," in *Proc. IEEE Int. Conf. Pattern Recognit.*, Aug. 2014, pp. 2065–2070.

[5] Y. Liu and Z. Wang, "Dense SIFT for ghost-free multi-exposure fusion," *J. Vis. Commun. Image Represent.*, vol. 31, pp. 208–224, Aug. 2015.

[6] T.-H. Oh, J.-Y. Lee, Y.-W. Tai, and I. S. Kweon, "Robust high dynamic range imaging by rank minimization," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 6, pp. 1219–1232, Jun. 2015.

[7] C. Lee and E. Y. Lam, "Computationally efficient truncated nuclear norm minimization for high dynamic range imaging," *IEEE Trans. Image Process.*, vol. 25, no. 9, pp. 4145–4157, Sep. 2016.

[8] N. K. Kalantari and R. Ramamoorthi, "Deep high dynamic range imaging of dynamic scenes," *ACM Trans. Graph.*, vol. 36, no. 4, pp. 144:1–144:12, Jul. 2017.

[9] S. Wu, J. Xu, Y.-W. Tai, and C.-K. Tang, "Deep high dynamic range imaging with large foreground motions," in *Proc. European Conf. Comput. Vis.*, Sep. 2018, pp. 120–135.

[10] Q. Yan, D. Gong, Q. Shi, A. van den Hengel, C. Shen, I. Reid, and Y. Zhang, "Attention-guided network for ghost-free high dynamic range imaging," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2019, pp. 1751–1760.

[11] B. Kim, H. Kim, K. Kim, S. Kim, and J. Kim, "Learning not to learn: Training deep neural networks with biased data," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2019, pp. 9004–9012.

[12] V. Monga, Y. Li, and Y. C. Eldar, "Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing," *arXiv preprint arXiv:1912.10557*, 2020.

[13] Y. Chen and T. Pock, "Trainable nonlinear reaction diffusion: A flexible framework for fast and effective image restoration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1256–1272, Jun. 2017.

[14] J. Zhang and B. Ghanem, "ISTA-Net: Interpretable optimization-inspired deep network for image compressive sensing," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 1828–1837.

[15] O. Solomon, R. Cohen, Y. Zhang, Y. Yang, Q. He, J. Luo, R. J. G. van Sloun, and Y. C. Eldar, "Deep unfolded robust PCA with application to clutter suppression in ultrasound," *IEEE Trans. Med. Imaging*, vol. 39, no. 4, pp. 1051–1063, Apr. 2020.

[16] C. Liu, J. Yuen, and A. Torralba, "SIFT Flow: Dense correspondence across scenes and its applications," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 5, pp. 978–994, May 2011.

[17] Y. Hu, D. Zhang, J. Ye, X. Li, and X. He, "Fast and accurate matrix completion via truncated nuclear norm regularization," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 9, pp. 2117–2130, Sep. 2013.

[18] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[19] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Jan. 2011.

[20] E. T. Hale, W. Yin, and Y. Zhang, "Fixed-point continuation for ℓ_1-minimization: Methodology and convergence," *SIAM J. Optim.*, vol. 19, no. 3, pp. 1107–1130, Oct. 2008.

[21] R. Mantiuk, K. Myszkowski, and H.-P. Seidel, "Lossy compression of high dynamic range images and video," in *Proc. SPIE*, Feb. 2006, vol. 6057, pp. 311–320.

[22] J. Froehlich, S. Grandinetti, B. Eberhardt, S. Walter, A. Schilling, and H. Brendel, "Creating cinematic wide gamut HDR-video for the evaluation of tone mapping operators and HDR-displays," in *Proc. SPIE*, Mar. 2014, vol. 9023, pp. 279–288.

[23] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, "Photographic tone reproduction for digital images," *ACM Trans. Graph.*, vol. 21, no. 3, pp. 267–276, Jul. 2002.

[24] T. O. Aydın, R. Mantiuk, and H.-P. Seidel, "Extending quality metrics to full luminance range images," in *Proc. SPIE*, Mar. 2008, vol. 6806, pp. 109–118.

[25] R. Mantiuk, K. J. Kim, A. G. Rempel, and W. Heidrich, "HDR-VDP-2: A calibrated visual metric for visibility and quality predictions in all luminance conditions," *ACM Trans. Graph.*, vol. 30, no. 4, pp. 40:1–40:14, Jul. 2011.